DOCUMENT RESUME

ED 029 663 LI 001 492

By-Avram, Henriette D.
The MARC Pilot Project. Final Report.
Library of Congress, Washington, D.C.
Spons Agency-Council on Library Resources, Inc., Washington, D.C.
Pub Date 68
Note-173p.
Available from-Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402 (GPO LC1.2iM18/2, $3.50)
EDRS Price MF-$0.75 HC Not Available from EDRS.
Descriptors-*Automation, *Cataloging, Catalogs, Computer Programs, Cost Effectiveness, Data Analysis, *Information Processing, Information Storage, *Libraries, *Library Cooperation, Library Programs, Library Technical Processes, Magnetic Tapes, Pilot Projects, Standards
Identifiers-*Machine-Readable Cataloging, MARC

The MARC Pilot Project was an experiment conducted by the Library of Congress, in cooperation with 16 participating libraries, to determine the feasibility of putting cataloging data into machine-readable from for distribution of magnetic tapes. MARC (Machine Readable Cataloging) records were distributed for some 16,000 titles in the pilot phase from November 1966 through June 1967 and for additional titles in the following fiscal year while the project was in transition to an operational stage. By June 30, 1968, approximately 50,000 records had been distributed to the participants for use in the production of book catalogs, catalog cards, book selection notices, and other library tools. This final report by the project's director contains a detailed description of the MARC pilot system, including the tape format, character sets, bibliographic codes, and input procedures. An analysis of the cost of production during the pilot period, as well as brief summaries of the computer programs used, are provided. The new MARC system which evolved from work carried out in the project and which is used for the MARC Distribution Service is also described. An appendix includes the reports and analyses of the pilot project written by the participating libraries. (LC Information Bull.)

ED029663

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

the MARC PILOT project

FINAL REPORT ON A PROJECT SPONSORED BY THE COUNCIL ON LIBRARY RESOURCES, INC.

LI001492

Prepared by
Henriette D. Avram
Information Systems Office
Project Director

LIBRARY OF CONGRESS • WASHINGTON : 1968

L. C. Card 68-67367

For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C., 20402 - Price $3.50

Foreword

The automation of any library activity poses unique problems and presents challenges not encountered in any other field. For the most part, this report, describing as it does a pioneering effort in cooperative library automation, presents problems and the solutions that were found for them. In order to guide those who will in the future automate other activities, it is fitting that the report cover these aspects in detail. The perceptive reader will, however, see beyond the problems and solutions to the contributions that many people have made to the success of this pilot project.

Behind the descriptions of computer programs and the data format lies a larger undescribed body of cooperation. The skills of computer specialists had to be coupled with those of librarians. The staff of the Library of Congress in the Information Systems Office had to be augmented by knowledge available among experts in other fields. Outside the Library of Congress, librarians and their technical colleagues who had developed or were developing systems contributed their advice and counsel at the beginning and in the course of the project. Participating libraries brought to bear not only their resources but, more important, their hopes and plans for the future.

It would be manifestly unfair to those who have cooperated in the project as well as to those who will use the results if this report were not as complete and as unbiased as possible. The Library of Congress, in its endeavor to serve the library community, therefore, feels that it is vital to progress to report all results, both good and bad, to serve as guides for future development of library automation. We need to know what went wrong as well as what went right with a process or technique in order to avoid repeating the mistakes of our predecessors.

The Library of Congress early recognized that the widespread application of computer technology to libraries could come about only if bibliographic data in machine-readable form could be distributed with precision and at reasonable cost. The Council on Library Resources, Inc., shared this recognition and granted funds for a beginning. The results have encouraged the Library to continue on a course which has now progressed beyond the critical pilot period. The regular distribution on a subscription basis of bibliographic data in machine-readable form, utilizing the refined and expanded format called MARC II, will begin in 1968. This service, which will augment the Library's distribution of printed catalog cards, should facilitate the development of automation throughout the entire library community.

L. QUINCY MUMFORD

Librarian of Congress

June 1968

Acknowledgments

The purpose of this report is to give libraries and persons concerned with library problems an account of the MARC (*MA*chine-*R*eadable *C*ataloging) Pilot Project and its continuing operation, the MARC Distribution Service. The report should provide a basis for understanding and judging the MARC project so that the library community, the Council on Library Resources, Inc., manufacturers, engineers, programmers, and others interested in the development of library automation can share in discussions and plans for the future.

The MARC project has been a sustained major undertaking carried out by a host of persons working at a variety of tasks. To avoid burdening the reader unduly, I must regretfully omit the names of many who made valuable contributions; to these individuals I must express my thanks collectively. I do wish however, to thank by name the individuals and groups mentioned below.

L. Quincy Mumford, Librarian of Congress, John G. Lorenz, Deputy Librarian of Congress, and Mrs. Elizabeth E. Hamer, Assistant Librarian, gave support and encouragement from the beginning. To Paul R. Reimers, Coordinator of Information Systems, I express appreciation for administrative and technical guidance, a rare blend in one individual. I wish to acknowledge the support of the Council on Library Resources, Inc., and Verner W. Clapp, its former president. For technical guidance and support of the project, I am indebted to colleagues at the Library of Congress: staff members of the Processing Department, Reference Department, Data Processing Office, Information Office, and Publications Office. To the staff members of Project MARC, I offer my warmest thanks. Their loyalty and dedication to the task has been and continues to be unwavering. I particularly want to recognize the work of Kay Guiles, who shared the responsibility of the implementation of the MARC Pilot System. In addition, I am indebted to him for all his efforts and patience during my indoctrination into the intricacies of bibliography and library procedures. Mrs. Lucia Rather and John Knapp provided valuable assistance in editing and preparing the report for publication. Peter Simmons was principally responsible for the cost analysis section of the report. He synthesized the work of Programming Services, Inc., and added a substantial portion by elaborating on the cost accounting still in operation at the Library. Other members of the MARC staff gave their valuable time to make this report as current as possible. Last but not least, I owe a debt of gratitude to MARC editors and paper tape typists. A large share of the success of MARC is theirs.

Our partners and our critics, the participants, deserve many thanks for guidance, encouragement, and constructive criticism. Special appreciation goes to Hillis Griffin, Argonne National Laboratory, for his untiring efforts in duplicating MARC tapes and listings for the participants. At the Library of Congress we are prone to forget that he is not a member of our staff. I am also indebted to United Aircraft Corporate Systems Center and to Programming Services, Inc., for their part in the MARC pilot experiment.

Thanks are especially due to Kathleen Bowman, Information Systems Office, Library of Congress. Without her ability to bring order out of chaos and her untiring effort at the typewriter in spite of constant revisions, this report would never have been completed.

Most of all, I am grateful to be a part of this exciting era in the history of libraries. For that, thanks to all librarians everywhere for the warm welcome and tremendous support I have received.

Henriette D. Avram

Table of Contents

Introduction

I Background and History of the MARC Pilot Project 3
II Objectives and Constraints of the MARC Pilot Project 9
III MARC Tape Formats 11
IV System Description 19
V Support Programs for the Participating Libraries 47
VI Codes Developed for Language, Publisher, and Place of Publication 53
VII Character Sets 59
VIII Cost Models 67
IX Evaluation of MARC I and Comparison with MARC II 77
X Potential Uses of Machine-Readable Data 83
XI Bibliography 85

Appendixes

A. Reports of Participants

Introduction 89
Argonne National Laboratory 91
Georgia Institute of Technology 95
Harvard University 103
Illinois State Library 113
Indiana University 115
Montgomery County Public Schools 119
Nassau (County) Library System 121
National Agricultural Library 123
Redstone Scientific Information Center 125
Rice University 127
University of California at Los Angeles 131
University of Chicago 135
University of Florida 137
University of Missouri 141
University of Toronto 143
Washington State Library 149
Yale University 167

B. Special Studies

Monograph Acquisition Patterns in Selected Libraries 175
Book-Oriented SDI Service Provided for 40 Faculty 179

Introduction

The MARC project has progressed from a pilot to test the feasibility of a distribution service of centrally produced machine-readable cataloging data to a full-scale operational system in the design stages in two short years. The library community, both here and abroad, has accepted MARC and recognizes its potential for the future. The single most significant result of MARC has been the impetus to set standards. There is no doubt that eventually standards would have been designed for machine-readable bibliographic records, character sets, and codes for place and language. MARC accelerated standardization and still more important, the standards are being set and agreed to by a large segment of the library community. The cooperation among the producers and users of bibliographic description has been a rewarding experience.

The extension of the MARC pilot into an operational system will have far-reaching implications. The much quoted article of Vannevar Bush[1] is most appropriately used in this context: "The difficulty seems to be, (sic) not so much that we publish unduly in view of the extent and variety of present-day interests, but rather that publication has been extended far beyond our present ability to make real use of the record. The summation of human experience is being expanded at a prodigious rate, and the means we use for threading through the consequent maze to the momentarily important item is the same as was used in the days of square-rigged ships." Libraries are on the brink of automation. It is to be hoped that the computer can be exploited as a tool in libraries so that they can better solve the needs of scholarship, science, and technology. MARC is a first step in this direction.

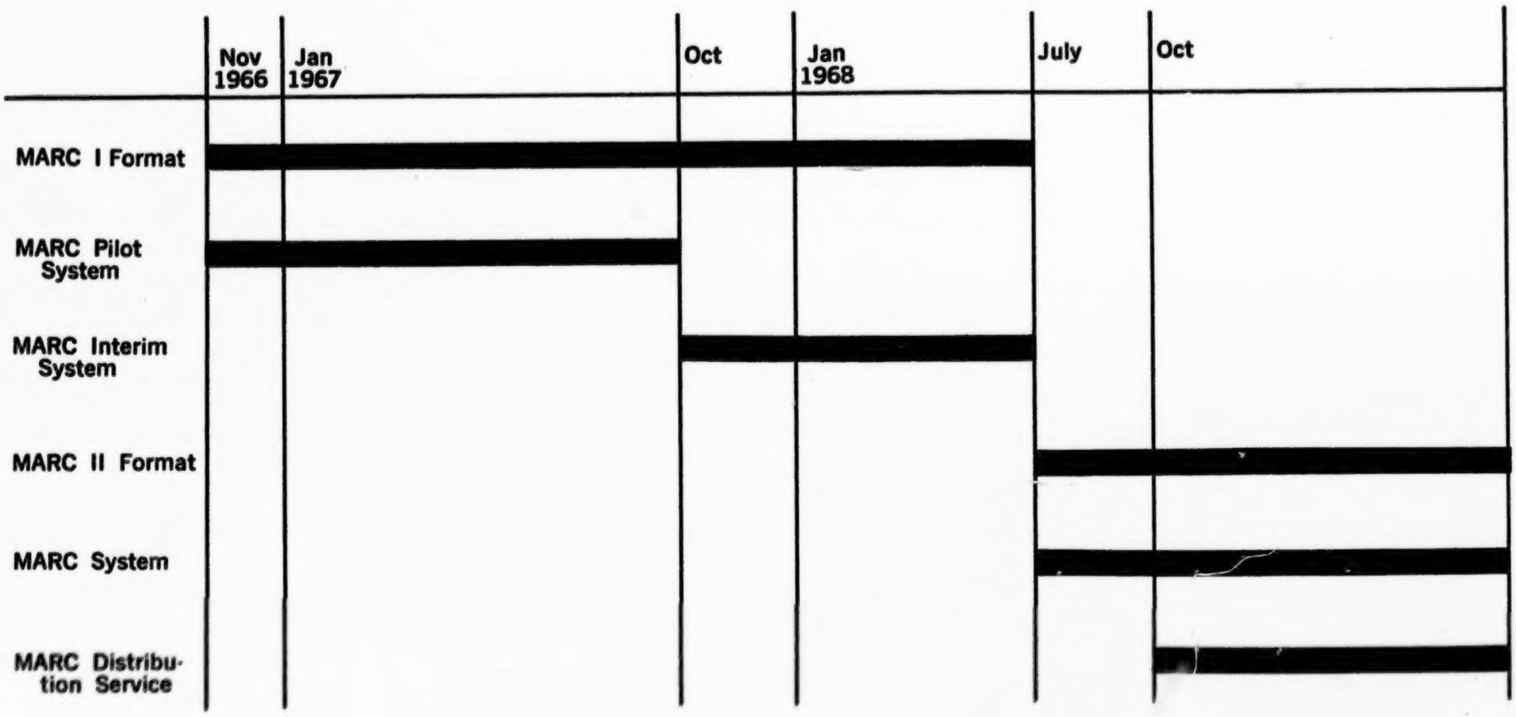
One of the most important products of an experimental project is the final report. The final report on the MARC Pilot Project has been delayed until now with a purpose in mind. It was clear that it would not suffice to describe the pilot project as an end in itself. The Library of Congress felt an obligation to the Council on Library Resources, Inc., and the library community to relate not only the results of the pilot but the lessons learned that led to the development of new ideas and finally resulted in new approaches. Learning, however, is a never-ending process. The MARC system will be evolutionary as needs are discovered in the Library of Congress and the scope of the system is expanded to satisfy these needs. The problem lies not in the discovery of things to do but rather in the ability to keep pace with a rapidly changing, dynamic library system. So the time has come to attempt to weave together in an orderly fashion results and accomplishments, knowing that tomorrow new ideas and fresh talent will make this publication out of date.

The decision to extend this report beyond the pilot project introduces problems of organization and definition. How best to describe the evolution of the MARC system and the interrelationships of the component parts to the whole has been a constant problem.

Let us consider the series of events as they occurred at the Library of Congress and sketch a framework with terms of references to be used as a guide to the organization and content of this report.

The MARC Pilot System, which included the first MARC format (hereafter referred to as the MARC I format), was in use at the Library of Congress from November 1966 until October 1967. At that time a new system was put into operation. This was called the MARC Interim System because work was already in progress at the Library to redesign the MARC I format. The new format would be called the MARC II format. The designers recognized that some part of the MARC Interim System would have to be modified to process the MARC II format, but it was considered desirable to use this experience to further refine the system. The MARC Interim System, which was in operation at the Library of Congress processing the bibliographic records in MARC I format, through June 1968, will be succeeded in July, when the Library will shift over to a period of testing and practice on the MARC II format and the modified programs and procedures necessary to produce the MARC II tapes. In October 1968, the MARC Distribution Service will make machine-readable cataloging data available to the library community. The bibliographic records will be in MARC II format; the operating system will be the MARC System.

For the purpose of this report, the operating system in use before October 1967 will be called the MARC Pilot System; the operating system between October 1967 and June 1968 will be the MARC Interim System. The system in use after June 1968 will be the MARC System. The following schematic should further amplify the above description:



Background and History of the MARC Pilot Project

The First Conference on Machine-Readable Catalog Copy

In 1964, the Council on Library Resources, Inc., (CLR) awarded a contract to Inforonics, Inc., for a study of the possible methods of converting the information on Library of Congress catalog cards to machine-readable form for the purpose of printing bibliographical products by computer. The result was the report, *The Recording of Library of Congress Bibliographical Data in Machine Form*,[2] submitted to CLR on November 23, 1964. To consider this report, a conference on machine-readable catalog copy was held on January 11, 1965, at the Library of Congress, under the sponsorship of the Library, the Committee on Automation of the Association of Research Libraries (ARL), and CLR. The First Conference on Machine-Readable Catalog Copy was also intended to provide a forum for discussion of related problems in different types of libraries. Participants in the conference included representatives from the Library of Congress as well as from universities, research agencies, Government agencies, and private industry.

The basis for the discussions was the format for a machine-readable catalog record and the bibliographical elements comprising this record as described in the Inforonics report. Other related topics included methods of distributing machine-readable catalog data to libraries, the future of card and book catalogs, and user-console dialogues. The conclusions reached by the conferees can be summed up as follows:

1. Early availability from the Library of Congress, by subscription, of machine-readable bibliographical data for current materials, as a byproduct of LC's cataloging operations, is desirable and will help libraries as they approach automated systems.
2. The Library recognizes the need for current cataloging data in machine-readable form, and it is seeking to identify the questions involved and to answer them. This will be done by expanding its automation program through increased staffing of its Information Systems Office and coordination of its various divisions.
3. The Library should probably include in its machine-readable record all data in the printed catalog card plus additional information. Most conferees favored coding as much data as possible to assure maximum future retrieval. It would seem desirable for the Library to go ahead with its own needs, other libraries using what they want from the LC machine-readable record.
4. The machine-readable record would be used for a variety of bibliographical products, such as card catalogs, book catalogs, bibliographies, acquisition lists, etc.
5. Agreement on data elements to be encoded is desirable, and the design of a machine record by the Library of Congress is probably the best means of standardization.

Three LC staff members were delegated to analyze cataloging data from a machine processing viewpoint. The results of this study were issued in June 1965 as *A Proposed Format for a Standardized Machine-Readable Catalog Record* (ISS Planning Memorandum No. 3).[3] This report suggests the contents of a machine-readable record, the manner of representing data, and the concept of fixed and variable fields applied to cataloging data as represented in machine-readable form.

The Second Conference

The proposed format for a machine-readable record was reviewed by over 150 Library of Congress staff members following publication, and their comments were issued as a supplement to the report. Comments were also elicited from libraries with ongoing automation projects and from major library interest groups. The memorandum was discussed at the October 1965 meeting of the Committee on Library Automation; it was distributed to the Automation Task Force of the Federal Library Committee; and it was the subject of a meeting of the Automation Committee of the Association of Research Libraries.

A meeting that would bring together representatives from different types of libraries was deemed beneficial both in identifying the common core of cataloging data that would serve them all and the supplementary data that would serve special requirements.

Accordingly, a second Conference on Machine-Readable Catalog Copy was planned by the Library of Congress. The Council on Library Resources agreed to support the conference and the meeting was held at the Library of Congress on November 22, 1965.

The fields to be represented in machine-readable form provided the core of the discussions, and possible magnetic tape formats containing the fixed and variable field data were explored. The discussions and the great interest shown by the participants in the Library's efforts strengthened the belief that the Library of Congress should begin an experimental magnetic tape distribution service as soon as possible.

The Initiation of the MARC Pilot Project

As a result of the second conference, the Library of Congress was encouraged to become the focal point for the exploration of the problems involved in the recording of bibliographic data in machine-readable form. In so doing it would draw heavily on the experience of the National Library of Medicine, University of Toronto, Florida Atlantic University, Yale University, and other agencies that already had projects underway. The Library had asked in June 1965 for a grant from the Council on Library Resources to explore the possibility of a distribution service. It now seemed feasible to augment the initial request to permit a more ambitious project, and a revised request was submitted. Funds were sought to support developmental work necessary to convert cataloging data into machine-readable form and distribute these data to a selected segment of the library community. The operation of the service was to be funded by the Library.

In December 1965 the Library received a grant of $130,000 to initiate a pilot project. The Council grant supported contracts to develop the required procedures for data conversion and the computer programs for input, file maintenance, distribution, etc. During the project, magnetic tapes would be mailed each week to a selected group of libraries. The project was called the MARC (for *MA*chine-*R*eadable *C*ataloging) Pilot Project.

Planning for the MARC Pilot Project began immediately. In February 1966 a contract was let to the United Aircraft Corporate Systems Center Division of United Aircraft Corporation to provide the necessary support in the design and development of the required system. The work included the development of computer programs for processing cataloging data at the Library of Congress and for printing the bibliographic record for the participants. It also included operating procedures for collecting, editing, transcribing, and distributing these data. Another contract was signed with Programming Services, Inc., to assist in the evaluation of the project and to design a cost model. Representatives of this organization made time studies, analyzed records containing schedules and work accomplished, studied feedback from participating libraries, and investigated data processing equipment and operations for means to improve the MARC distribution.

From 40 libraries that had expressed willingness to participate in the project, 16 participants were selected. The selection was based on several factors, including type of library (special, Government, State, university, public, and school); geographical location; availability of personnel, equipment, and funds; proposed use of the MARC data; and expressed willingness to evaluate the utility of the data and prepare reports. The original MARC participants were:

Argonne National Laboratory, Argonne, Illinois

Georgia Institute of Technology, Atlanta, Georgia

Harvard University, Cambridge, Massachusetts

Indiana University, Bloomington, Indiana

Montgomery County Public Schools, Rockville, Maryland

Nassau (County) Library System, Hempstead, L.I., New York

National Agricultural Library, Washington, D.C.

Redstone Scientific Information Center, Redstone Arsenal, Alabama

Rice University, Houston, Texas
University of California Institute of Library Research, Los Angeles, California
University of Chicago, Chicago, Illinois
University of Florida, Gainesville, Florida
University of Missouri, Columbus, Missouri
University of Toronto, Toronto, Ont., Canada
Washington State Library, Olympia, Washington
Yale University, New Haven, Connecticut

The Third Conference

With the support of the Council on Library Resources, a Third Conference on Machine-Readable Catalog Copy was held at the Library of Congress on February 25, 1966. This meeting brought together representatives of the participating libraries, Library of Congress staff members, and United Aircraft personnel assigned to the MARC project. The concept of the project was described and the roles of the participants, the Library of Congress, and the contractor were discussed. This meeting was the official opening of the MARC Pilot Project. The topics covered included:

1. The concepts, objectives, schedules, functions, and requirements of the experiment.
2. The operation at the Library of Congress.
3. The MARC I format for the data, including variable and fixed fields.
4. Weekly materials sent to participants by the Library of Congress.
5. The responsibilities of the participants.
6. Uses participants planned to make of MARC records.
7. Reports from participating libraries to LC.

The Information Systems Office announced that the MARC I format would be completed by April 1966 and that distribution of tapes to the participants would begin by September of that year.

Implementation of the Project

This schedule, however, could not be kept. The distribution system planned for September actually began with the mailing of one test tape in October followed by the regular weekly distribution in November.

The first tapes included both catalog and cross-reference records. It soon became evident that although the cross-reference information was valuable, it was not as useful as had been anticipated because the cross-reference records were not linked to the associated catalog records. Distribution of cross-reference information was discontinued in May 1967 with the understanding that the problem would receive further investigation.

The pilot project was scheduled to end on June 30, 1967. Encouraged, however, by the enthusiasm expressed by the library community, at the Midwinter Conference of the American Library Association in January 1967, the Library of Congress announced its intent to extend the project beyond June.

In March 1967 the MARC staff began a formal evaluation of the MARC I format. (See chapter 9 for format discussion.) A preliminary MARC II format design was presented to the participants for their comments during the MARC meeting held at the June 1967 ALA conference in San Francisco. At this meeting the Library of Congress also announced that a full-scale operational MARC Distribution Service would begin in 1968 and that the project would continue during the next fiscal year (June 1967-June 1968) while the new service was planned and implemented.

During the interim period, the Library felt that the distribution service could be extended to a few more libraries that had expressed interest in participation. Letters of invitation were extended and four libraries were chosen in January 1968 as participants until the time that the pilot distribution would end. The libraries selected were California State Library, Illinois State Library, Cornell University Library, and SUNY Biomedical Communications Network, Syracuse, N.Y.

A Fourth Conference on Machine-Readable Catalog Copy, supported by the Council on Library Resources, was held at the Library of Congress, December 4, 1967. The purpose of this meeting was to discuss the MARC II format and a proposed character set for bibilographic data. In addition, the feasibility of sharing computer program specifications in the library community was explored.

The significance of the meeting was that the Library and the participants, through their experiences in the MARC Pilot Project, were now cooperatively involved in the development of the operational system which would have wide implications for the entire library community and the future of automated library systems.

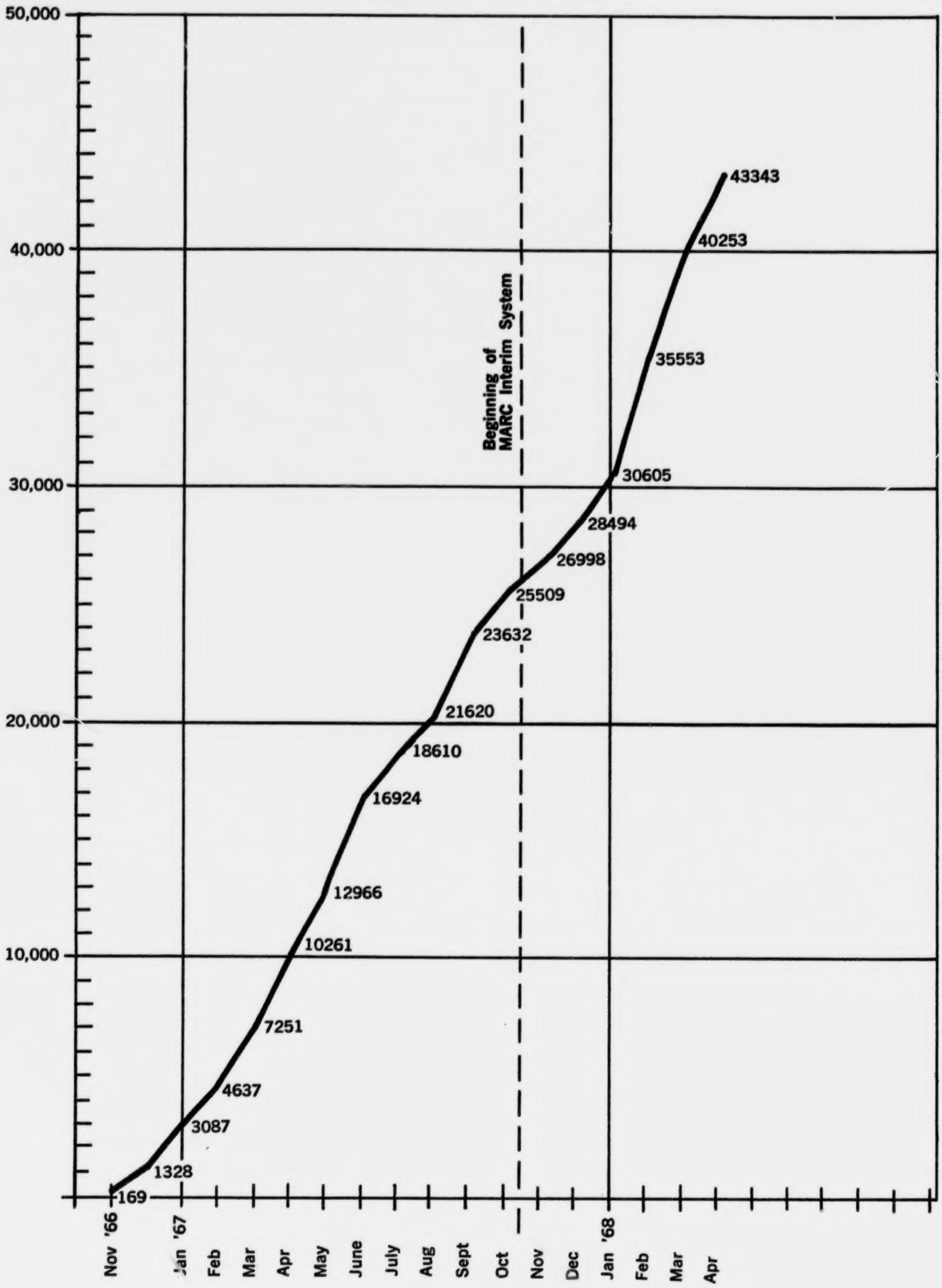


FIGURE 1.—*Cumulative MARC Production*

By May 1968 the MARC tapes included over 44,000 titles in machine-readable form. It was anticipated that this number would reach 50,000 by the time the project ended in June 1968. Figure I shows cumulative production during the project.

TABLE I.—*MARC Records by Major Library of Congress Classification up to April 5, 1968*

CLASS	TOTAL	CLASS	TOTAL
A	140	M	458
B	3384	N	1536
C	494	P	10,562
D	3105	Q	3186
E	1039	R	1485
F	1086	S	718
G	1275	T	2856
H	4616	U	298
J	1163	V	164
K	470	Z	757
L	1096	No Class	988
Total			40,876

Objectives and Constraints of the MARC Pilot Project

Objectives

The MARC Pilot Project was an experiment to determine the feasibility of centrally producing a standardized machine-readable record for application by local installations to serve their specific requirements. To accomplish this, the project was to provide machine-readable cataloging data to a representative segment of the library community and then to assess the utility of the bibliographic information supplied. The participating libraries, within the scope of their individual requirements, would utilize the MARC record in producing typical library records, such as catalog cards, book catalogs, bibliographies, accession lists, etc.

It was anticipated that there would be numerous side benefits from the experiment. Information would be provided of immediate use for the system development study of the central bibliographic operation at the Library of Congress. The data base obtained would be useful for research in areas such as book catalog production, file organization, retrieval modes, and man-machine dialogues. Data would be obtained to aid the Library of Congress and the library community to plan more ambitious projects of greater scale and sophistication. Any full-scale operational MARC system would benefit from the analysis of pilot input procedures, worksheets, costs, methods of distribution, character sets, etc. The participants would report the results of their activities and this information would provide guidelines for the use of MARC data in local institutions.

Of major significance would be the evaluation of the machine-readable format by both the Library of Congress and the participants.

Constraints

There was great interest in determining, as rapidly as possible and with reasonable expenditure of resources, the feasibility and utility of centrally prepared machine-readable cataloging data. In the MARC Pilot Project, budget and time constraints influenced design considerations. Among these were:

1. Project Facilities. The implementation of the project required a facility for the central preparation of machine-readable catalog records. A major change to the computer configuration would be expensive and could not be justified. The decision was made to augment the existing equipment at the Library of Congress with a paper-tape reader, additional magnetic-tape drives, selector channels, and a specially designed print train. This decision influenced the design of computer programs.
2. Mode of Data Collection. In view of the time constraint and the experimental nature of the project, it was deemed inadvisable to disturb the existing internal operations of the Processing Department of the Library of Congress. A MARC System Production Group was organized within the Information Systems Office to receive cataloging data and prepare it for conversion to machine-readable form, without interfering with normal cataloging procedures. The manuscript card prepared by the Descriptive Cataloging, Shared Cataloging, and Subject Cataloging Divisions of the Library was reproduced on a preprinted input worksheet which became the source data for MARC. The design of the format, the worksheets, and the editing procedures were all influenced by the use of source data in this form.
3. Data Base. Since it was more difficult to find personnel with expertise in foreign languages, the decision was made to limit the pilot project to current English language monographs. During the planning stages of MARC, the English language cataloging output was computed to be 600 titles per week. During the life of the project, however, Title II-C of the Higher Education Act became a reality, and the number of English language monographs cataloged exceeded 1,200 per week.

4. Number of Participants. At the onset of the project, there were many unknown variables. Projected costs to be met by funds appropriated to the Library of Congress placed a limitation on the number of participants at 16.

MARC Tape Formats

The MARC tape distributed to the participants contained four separate files of information:

File 1: Machine-Readable Catalog (MARC) Record

File 2: Machine-Readable Author/Title Record

File 3: Machine-Readable Subject Cross-Reference Tracing Record

File 4: Machine-Readable Descriptive Cross-Reference Tracing Record

Four files were always recorded on the tape; either full data files or, if a file was not used, an end-of-file mark was written on the tape for that file.

The MARC tape was produced in both 7-level and 9-level form. Nine-track tapes were written with odd parity and a tape density of 800 cpi; 7-track tapes had even parity and density of 556 cpi. Because a standardized magnetic tape format was desired for all MARC tapes, character positions 1 through 8 of each record contained block and record length information for the IBM System/360 users and blank characters (block and record length irrelevant) for the IBM 1401 users.

Files 1 and 2 contained cumulative records for the period of use of each tape. Files 3 and 4, the cross-reference tracing records, were accumulated for two weeks only.

File 1—Machine-Readable Catalog (MARC) Record

The cataloging information in the MARC record was divided into two sections: the fixed fields and the variable length fields.

FIXED FIELDS

Fixed fields contained information about a catalog record which could be coded in a predetermined number of characters. That is, the number of characters in each field was the same from record to record, and each coded symbol was located in the same position in every record. The fixed fields used in the MARC record were as follows:

1. Block Length. This four-character field was always equal to the record length plus four. The first two characters specified the length of the block in 16-bit binary form; the next two characters were blank. The maximum block length was 1,350 characters.

2. Record Length. This four-character field contained the record length. The first two characters specified the length of the record in 16-bit binary form, and the next two characters were blank. The maximum record length was 1,346 characters.

3. Library of Congress Catalog Card Number. This 11-character field allowed for three leading alpha characters and eight numerics. Alpha characters were used to represent any prefix to the LC catalog card number. The prefix number was left-justified with the remaining spaces filled with blanks; if there was no prefix, three blanks appeared before the numeric card number. The numeric part of the LC card number was an eight-digit number; the first two digits were a date and the last six digits were an identification number. The date always appeared in character positions four and five of the 11-character field, while the identification number was right-justified with all leading spaces filled with zeros.

Example:

Number on Printed Card	*Number on Magnetic Tape*
A66-11	Ab̸b̸66ØØØØ11
AB66-111	ABb̸66ØØØ111
66-1	b̸b̸b̸66ØØØØØ1

(In these examples, "b̸" stands for blank and represents one character position. "Ø" indicates zero; the slash differentiates between zero (Ø) and the letter O.).

4. Supplement Number. Since supplements, indexes, and other dashed-on entries were carried as independent records in the MARC Pilot Project, yet had the same LC catalog card number as the original work, it was necessary to provide a field of one numeric character to indicate a supplement and specify its number. If the record was not a supplement, this field was blank.

TABLE 2.—*List of Fixed Fields*

Fixed Field	Character Position in Record
1. Block Length	1–4
2. Record Length	5–8
3. Library of Congress Catalog Card Number	9–19
4. Supplement Number	20
5. Type of Main Entry	21
6. Form of Work	22
7. Bibliography Indicator	23
8. Illustration Indicator	24
9. Map Indicator	25
10. Conference or Meeting Indicator	26
11. Juvenile Indicator	27
12. Language Indicator	28
13. Language 1	29–32
14. Language 2	33–36
15. Type of Publication Date	37
16. Date 1	38–41
17. Date 2	42–45
18. Place of Publication	46–49
19. Publisher	50–53
Unassigned Field (Blanks)*	54–58
20. Height of Volume	59–60
21. Types of Secondary Entries	61–68
22. Series Indicator	69
23. Local Use	70–103
24. Control Indicator	104
25. Length of Record	105–108

* This is a field of five characters reserved for LC use.

5. Type of Main Entry. One alpha character identified the type of main entry in accordance with the *ALA Cataloging Rules for Author and Title Entries* as follows:

A—Personal Author
B—Government Body
C—Society or Institution
D—Religious Society or Institution
E—Miscellaneous Corporate Body
F—Uniform
G—Title

6. Form of the Work. In the MARC Pilot Project, only two forms were identified.

M—Monograph
S—Serial

7. Bibliography Indicator. If the work contained a bibliography or was itself a bibliography, this fixed field contained an X; otherwise the field was blank.

8. Illustration Indicator. If the work contained any type of illustrations other than maps, this fixed field contained an X; otherwise, the field was blank.

9. Map Indicator. If the work contained maps, this fixed field contained an X; otherwise, the field was blank.

10. Conference or Meeting Indicator. If the work contained the proceedings or the report of a conference, meeting, or symposium, this fixed field contained an X. Otherwise, the field was blank.

11. Juvenile Indicator. If the work was for juveniles (as indicated, for example, by the subject heading and/or the classification number) this field contained an X. Otherwise, the field was blank.

All records in the Annotated Card (AC) Program contained an X in this field. Since the same basic MARC record could appear in two different formats (that of the annotated card and that of the standard LC catalog card), it was sometimes represented twice on the MARC tape. Annotated cards did not, however, have the same LC catalog card number as their counterparts in the regular card program.

12. Language Indicator.

13. Language 1.

14. Language 2.

The alpha character that appeared in the language indicator described the use of languages in the work and determined the content of the two language fields (which had either three or four alpha characters). If language 1 or 2 had only three characters, the code was left-justified with the fourth character position in the field a blank. The language indicator had five alternate entries:

S—The work contained only one language. The language was given in the language 1 field and the language 2 field was blank.

T—The work was a translation. Language 1 indicated the language of publication; the language 2 field indicated the language in which the work was originally written or multilingual, as in the case of anthologies.

M—The work contained more than one language. The principal language was given in the language 1 field; if only two languages were used, the second was given in the language 2 field. If there were more than two languages, either multilingual was given as the second language or the predominant language was given in three characters with the fourth character indicating multilingual.

D—The work was a dictionary of more than one language. If only two languages were used, these were given in the language fields. If more than two languages were used, the language 1 field contained the principal language, and in the language 2 field either multilingual was given as the second language or the second predominant language was coded in three characters with the fourth character indicating multilingual.

G—The work was a grammar or reader of the type used in language courses. The native language was indicated in the language 1 field; the language being studied was given in the language 2 field. (See Chapter 6 on language codes).

15. Type of Publication Date.

16. Date 1.

17. Date 2.

The contents of the two (numeric) dates of publication fields (date 1 and date 2) were determined by the alpha code that appeared in the type of publication date field as follows:

S—The date of publication consisted of a known date or a probable date that could be represented by four digits. The date was given in the date 1 field. The date 2 field contained the date of copyright if it appeared in the imprint statement in addition to the date of publication.

R—The work was a reproduction (such as a reprint or facsimile). The publication date of the reproduction was given in the date 1 field. The date 2 field contained the date of original publication.

N—The date of publication was not known. Both date fields were blank.

M—The date of publication consisted of multiple dates. The initial date was given in the date 1 field. When the terminal date was known, it was given in the date 2 field; otherwise, the date 2 field was set to the year 9999 to indicate an open-ended situation.

Q—One or more of the digits in the imprint date was missing. The dates entered in the date 1 and date 2 fields were those which bracketed the time period indicated by the incomplete imprint date, as shown in these examples:

	Key	*Date 1*	*Date 2*
18—	Q	1800	1899
189-	Q	1890	1899

18. Place of Publication. This field contained four alpha characters representing a place of publication. (See Chapter 6 for place codes.)

19. Publisher. This code had two, three, or four alpha characters. If the code contained only two or three characters, it was left-justified within the field and the remaining character positions in the field were blank. (See Chapter 6 for publisher codes.)

20. Height of the Volume. Two numeric characters represented the height of the volume in centimeters. If the height of the volume was a fractional number, the next higher whole number was entered; for example, 25½ centimeters was entered as 26 centimeters.

21. Types of Secondary Entries. An X in a character position in this field indicated the presence of a specific type of secondary entry traced in the record; a blank indicated the absence of the entry.

Position	Indication
61	Name of at least one personal author
62	Name of at least one government body
63	Name of at least one society or institution
64	Name of at least one religious society or institution
65	Name of at least one miscellaneous corporate body
66	At least one uniform heading
67	At least one title secondary entry
68	At least one subject heading

The entries in this field were set by the computer programs in the processing of variable-field information.

22. Series Indicator. If the work was a part of a series, this field contained an X; otherwise, the field was blank. This indicator was set by the computer programs whenever a series was encountered in the processing of variable field information.

23. Local Use. Character positions 70 to 103 were set aside for use by the participating libraries.

24. Control Indicator. One alpha character was used to indicate the current status of the record on the MARC tape:

N—Record is new this week
O—Record was new last week
R—Record has been revised this week
Blank—Record is at least two weeks old

25. Length of Record. This field of four numeric characters indicated the total number of characters in the record, beginning with the first character of the LC catalog card number and ending with the last character of the last variable field.

VARIABLE FIELDS

In each variable field, the first three character positions gave the length of the field. The characters in positions 4 to 6 contained the identifying tag. The first two characters of the tag are shown in Table 3. The third character of the tag was used only in certain variable fields; all other fields had this position blank. The following descriptions of the variable fields indicate the use of a third character where relevant. They are arranged in the order in which the variable fields appear in the magnetic tape record, whereas in Table 3, they are listed in numerical sequence by tag number.

TABLE 3.—*List of Variable Fields*

Description	Tag Number
Main Entry	10
Conventional or Filing Title	15
Title Statement	20
Edition Statement	25
Imprint Statement	30
Collation Statement	40
Series Note (to be an added entry and to be traced in exactly the same form as in the Series Note)	50
Series Note (to be an added entry but not to be traced in the form in the Series Note/or not to be an added entry)	51
Notes	60
Subject Tracing	70
Personal Author Tracing	71
Corporate Author Tracing	72
Uniform Tracing	73
Title Tracing	74
Series Tracing	75
Copy Statement	80
National Bibliography Number	83
Library of Congress Call Number	90
Dewey Decimal Classification Number	92
Library of Congress Catalog Card Number	94

1. Library of Congress Call Number (tag 90). This field contained a complete LC call number or class number or did not appear. The presence of a class number only indicated that the Library of Congress had cataloged the publication but had not planned to add it to its collections. In the MARC Pilot Project these class numbers were not enclosed within brackets as they are on the LC typeset cards, but instead were followed by the letters NLC. These letters, defined as "not in LC," were separated from the class number by two spaces.

The absence of a LC call number indicated that the publication had been assigned to the Law Library of the Library of Congress.

2. Dewey Decimal Number (tag 92). This field contained the complete Dewey Decimal number. When these numbers were prefixed by the letter j, they appeared in the MARC record with the j following the Dewey Decimal number and separated from it by two spaces and enclosed within parentheses. The j was used to designate works for juveniles and the numbers were assigned from the 9th edition (abridged) of the *Dewey Decimal Classification and Relative Index.* Works in the Annotated Card Program sometimes had an E (easy book) or Fic (fiction) entered in place of the Dewey Decimal number when appropriate.

3. Main Entry (tag 10). The format of the main entry was that specified in the *ALA Cataloging Rules for Author and Title Entries,* with the following exceptions for personal names:

Titles followed the forename rather than the surname, e.g., Scott, Walter, Sir, bart., 1771-1882.

Date modifiers (b., d., fl., etc.) followed rather than preceded dates, e.g., 1281, fl.

All initials were closed up except those representing personal names, e.g., IBM, U.S., but Harris, A. E.

Tag 10 was not used for a title entry; tag 20, for title statement, was used instead. The pound sign (#) was used as a delimiter in all names and headings used as a main entry.

Personal names followed a pattern of name, title, date, and relator. Title included all designations of rank, office, or nobility, or words or phrases associated with a name. Relator referred to those phrases describing the relationship between a name and a work (e.g., ed., tr., and comp.). The following combinations of these elements depict the use of the pound sign as a delimiter:

Name
Name, # Date#
Name, # Date#, Relator
Name, # Title
Name, # Title, Date#
Name, # Title, Date#, Relator
Name, # Title, Relator
Name, # Relator

After a name, the delimiter followed the punctuation. After a date, the delimiter followed the fourth character of the date.

Examples:

Churchill, Winston Leonard Spencer,
# Sir, 1874#-1965
Smith, John, # 1859#, d.

In the cases of main entries other than personal names, a delimiter was inserted following the last character that would normally appear in bold face type on an LC typeset card.

Examples:

U.S.# *Library of Congress.*
California Institute of Technology,
Pasadena.

4. Conventional or Filing Title (tag 15). The contents of this field reflected current Library of Congress filing procedures.

5. Title Statement (tag 20). This field contained the title and all subsequent information up to, but not including, the edition statement. The pound sign was used to define the end of the title and the end of the short title, if one existed.

6. Edition Statement (tag 25). This field contained the complete edition statement. The pound sign was used to separate edition information from the remainder of the statement.

7. Imprint Statement (tag 30). This field contained the complete imprint statement and the price of the work if it had been cataloged under the National Program for Acquisitions and Cataloging. The pound sign was used to separate place, publisher, date, and price. The pound sign followed the punctuation mark after the place and publisher; the presence of only two delimiters indicated the absence of a publisher in the imprint statement. (Place included n.p. and date included n.d.).

Examples:

Place, # Publisher, # Date#
Place, # Publisher, # Date# Price
Place, # Date# Price
Place, # Date#

8. Collation Statement (tag 40). This field contained the complete collation statement, i.e., paging, illustration, and size.

9. Series Notes (tag 50). Tag 50 indicated that the series was to be traced in exactly the same form as it appeared in the series note. When the series note consisted of an author and a title, a dollar sign used as a delimiter defined the end of the author element and the beginning of the title. Tag 50 was repeated as often as necessary.

10. Series Note (tag 51). Tag 51 was used for each series that was not to be traced in the same form as it appeared in the series note or that was not to be traced at all. Tag 51 was repeated as often as necessary.

The first series note to appear in the record, whether designated by a 50 or a 51 tag, was the one placed in parentheses following the collation statement. Fields tagged 50 or 51 could appear in any sequence (for example, 50, 51; 51, 50) or any combination for as many series notes as appeared in the record.

11. Notes (tag 60). The information contained in each note appeared as a separate variable field in the record. Tag 60 was repeated as often as necessary.

12. Subject Tracings (tag 70). Personal names used as subject tracings were formatted in the same pattern as they were when used as main entries; however, no pound sign delimiters were used. Tag 70 was repeated as often as necessary. For two-line tracings (such as an Author/Title tracing), the dollar sign delimiter defined the end of the first element for overprinting of the added entry.

13. Personal Name Tracings (tag 71). The format and the use of pound sign delimiters for the tracings followed the same rules as applied to personal name main entries. Tag 71 was repeated as often as necessary. For two-line tracings, the dollar sign defined the end of the first element for overprinting of the added entry.

14. Corporate Name Tracings (tag 72). The corporate name tracing, other than subject, was designated by a three-character tag. The first two characters of the tag were 72, and the third character specified one of four types of corporate names:

72B—Government Body
72C—Society or Institution
72D—Religious Society or Institution
72E—Miscellaneous Corporate Body

The pound sign delimiter was not used. Tag 72 was repeated as often as necessary. For two-line tracings, the dollar sign delimiter defined the end of the first element for overprinting of the added entry.

15. Uniform Heading Tracings (tag 73). The uniform heading tracing used as an added entry was designated by a tag 73. There were no pound signs used as delimiters, and tag 73 was repeated as often as necessary.

16. Title Tracings (tag 74). When a full title or a short title was to be traced in exactly the same form as the title statement, the tag 74 field contained only a T in the seventh character position, the first position of the variable content. When a title tracing was not the same as the full title or short title in the title statement (tag 20), the character T appeared in the seventh position of the tag 74 field and the title tracing followed in complete form. Tag 74 was repeated as often as necessary.

17. Series Tracings (tag 75). This field contained either an A or T in the seventh character position. An A indicated a series consisting of an author and a title; a T indicated a series tracing consisting of a title only.

If the content of the series tracing had been defined by a tag 50 series note, the tag 75 field contained only seven characters, with A or T as the seventh character. If the content of the series tracing had been defined by a tag 51 series note, the tag 75 field contained either A or T in the seventh position of the field, followed by the series added entry as it was to be traced. When the content of a series tracing consisted of an author and a title, the dollar sign defined the end of the author element and the beginning of the title. When tag 75 consisted of seven characters only, the dollar sign was carried in the tag 50 field. Tag 75 was repeated as often as necessary.

18. Copy Statement (tag 80). This field contained information used by the Library of Congress.

19. National Bibliography Number (tag 83). A national bibliography number is the item number of a title listed in a national bibliography. The third character of this tag for this field was either a one or a zero. A zero indicated that the national bibliography number could be accommodated by 15 characters or less. A 1 indicated that the number required more than 15 characters. This latter case usually occurred when the national bibliography numbers for a multivolume work were listed. In this situation the field contained the first national bibliography number listed on the printed LC catalog card; the entire series of numbers made up the first note tagged 60.

20. Library of Congress Catalog Card Number Suffix (tag 94). An LC catalog number with a date and number (66–1037) or with prefix, date, and number (AC66–1037) appeared in complete form (prefix and numerics) in the fixed field. However, when the LC catalog card number included a suffix (66–1037/CD), the suffix did not appear in the fixed field but appeared with a tag 94 in a variable field. A tag 94 field was present only when there was a suffix; it had a maximum length of three character positions.

File 2—Machine-Readable Author/Title Record

The records in the author/title file were of a fixed length of 134 characters, were unblocked, and were sorted alphabetically by author/title.

AUTHOR INFORMATION

The author information was extracted from the main entry (variable field tag 10) in the MARC record up to a maximum of 79 characters. When the author information on the MARC record exceeded 79 characters, character positions 77 to 79 were set to periods to signify that the complete author information was not given. Personal name information included only the data up to the second delimiter or the end of the tag 10 field, whichever occurred first. The first delimiter in the author information was deleted from the record.

TITLE INFORMATION

Sufficient title information was extracted from the title statement (variable field tag 20) in the MARC record to complete the 120-character author/title field, which had a minimum of 40 characters of title information. For a title main entry, only title information was present in the field. If the data in this field did not fill the entire 120-character positions, trailing blanks were added to complete the field.

AUTHOR/TITLE RECORD FORMAT

Character positions 1 through 11 contained the LC catalog card number allowing for three leading alpha characters and eight numerics. Position 12 contained the supplement number, when present, and position 13 was used for the type of main entry code. Character position 14 contained one of four control indicators:

Indicator	*Interpretation*
N	New entry this week
O	New entry last week
R	Record revised this week
ƀ	Record at least two weeks old

Character positions 15 through 134 contained the author/title information in two subfields of variable length, separated by the pound sign. The combined size of these two subfields, including the delimiter, could not exceed 120 characters. When the data required fewer than 120 characters, trailing blanks were added to fill the field.

Files 3 and 4—Machine-Readable Cross-Reference Tracing Records

Cross-reference tracing information consisted of information that allowed the generation of cross references relevant to a particular heading. A cross-reference tracing record consisted of a heading and the tracing of the cross reference(s) related to that heading. For example:

Gt. Brit. Commonwealth Immigrants Advisory
Council
x Gt. Brit. Home Dept. Commonwealth
Immigrants Advisory Council;
Gt. Brit. Immigrants Advisory Council

The cross-reference tracing information included in the MARC Pilot Project was that information generated in the cataloging of monographs which were incorporated into File 1. The source of information for subject headings and their related cross-references was the Subject Cataloging Division of the Library of Congress. The source of information for the headings and related cross-references established by descriptive catalogers was the Descriptive Cataloging and Shared Cataloging Divisions of the Library of Congress. The headings and cross references for which the two latter divisions were responsible will be referred to hereafter as "descriptive headings" and "descriptive cross references," respectively. Any mention of the Descriptive Cataloging Division should be taken to include the Shared Cataloging Division as well. The cross-reference tracing information from the three divisions was processed as it was received by the MARC System Production Group.

The formats for Files 3 and 4 were identical. They were variable in length and unblocked. File 3 contained Subject Cross-Reference Tracing Records and File 4 contained Descriptive Cross-Reference Tracing Records (see Table 4) arranged alphabetically by heading.

The tag of a subject heading field was 10, and of a descriptive heading, 50; the third character of the three-character heading tag subfield was normally blank. The reference fields were arranged numerically by tag. Each group of similar references was tagged only once; each reference within the field was separated by delimiters. The last reference in the group was not followed by a delimiter. Subject references were tagged with the number 20 for *sa* references, 30 for *x* references, and 40 for *xx* references. Descriptive references were tagged with the number 60 for *x* references and 70 for *x* (*see also*) references.

RECORD CODES

There were specific optional action indicators that were used in position 20 of a cross-reference tracing record to describe the type of information in the record or action that had been taken. The code used in position 19 (the third position of the heading tag) further defined the elements affected by the action indicator. These codes are shown in Table 4.

Position 20	Position 19	Description
N	Blank	New reference data
A	Blank	Added reference data
C	H	Heading and all related references canceled
C	F	Only reference fields canceled
R	F	Reference corrected

An N indicated that the cross-reference information in the record was new, that is, both the heading and the references were new. An A indicated that the references traced were new but that the heading had been previously established by the Library of Congress. There was no indication as to whether the heading to which a reference was being added had previously appeared in the MARC Pilot System, nor any notation of any references previously traced for this heading.

The letter C indicated the cancellation of either the heading or the references traced. When both the heading and one or more references were given, the specific references had been canceled. When the heading alone was given, the heading and all related references had been canceled. For each C indicator (in position 20 of the field), the third character (position 19) of the heading tag field contained an H or F indicator which specified the canceled fields: an H indicated that the heading and all related references had been canceled; an F indicated that only the reference fields given had been canceled.

An R indicated that an error had been made in the reference (in a reference field) of the record previously entered into the MARC Pilot System and the correct information was then given. The heading field contained the heading originally supplied; the reference field contained all previously supplied correct references and the correct form of the references previously supplied in an incorrect form. If there were other reference fields in the original record that were not affected by the revision, they did not appear in the revision record.

A revision of a heading was accomplished by a different method from that used for revisions of references. The heading was changed through a two step operation of "cancel" and "new."

TABLE 4.—*List of Fields in Cross-Reference Tracing Records*

Field	Character Positions on Record
Block Length	1–4
Record Length	5–8
*Record Length	9–12
Control Indicator	13
Heading Indicator	
Field length	14–16
Tag of the field	17–19
Action indicator	20
Pound sign delimiter	21
Heading	Variable
Cross-Reference Fields	
Field length	1–3 positions in cross-reference field
Tag of the field	4–6
Cross-reference field	Variable

---

* This four-character fixed field (numeric) indicated a total number of characters in the record beginning with its own first character and ending with the last character of the last variable field.

System Description

Activity within MARC begins with the receipt of a bibliographic record in the form of a reproduction of the manuscript card prepared by the Library of Congress catalogers. This manuscript card, used to produce the typeset Library of Congress catalog card, is reproduced on an input worksheet and becomes the source data for MARC. The information on this sheet is edited, punched on a paper tape typewriter, and input to the computer for transfer to magnetic tape. The data undergo two processing cycles, a daily and a weekly, and are output on a MARC master tape which is duplicated for distribution to libraries.

The processing of bibliographic data for the generation of the MARC record encompasses four major functions: data collection, data preparation, data transcription, and computer processing. The following sections describe these functions in the daily and weekly processing routines through which each MARC record flows during its transformation into an entry on the MARC master tape. Each section describes the system as it was designed for the MARC Pilot System, modified for the MARC Interim System, and projected for MARC System.

Data Collection

Data collection is the gathering together of the information to be processed for an application.

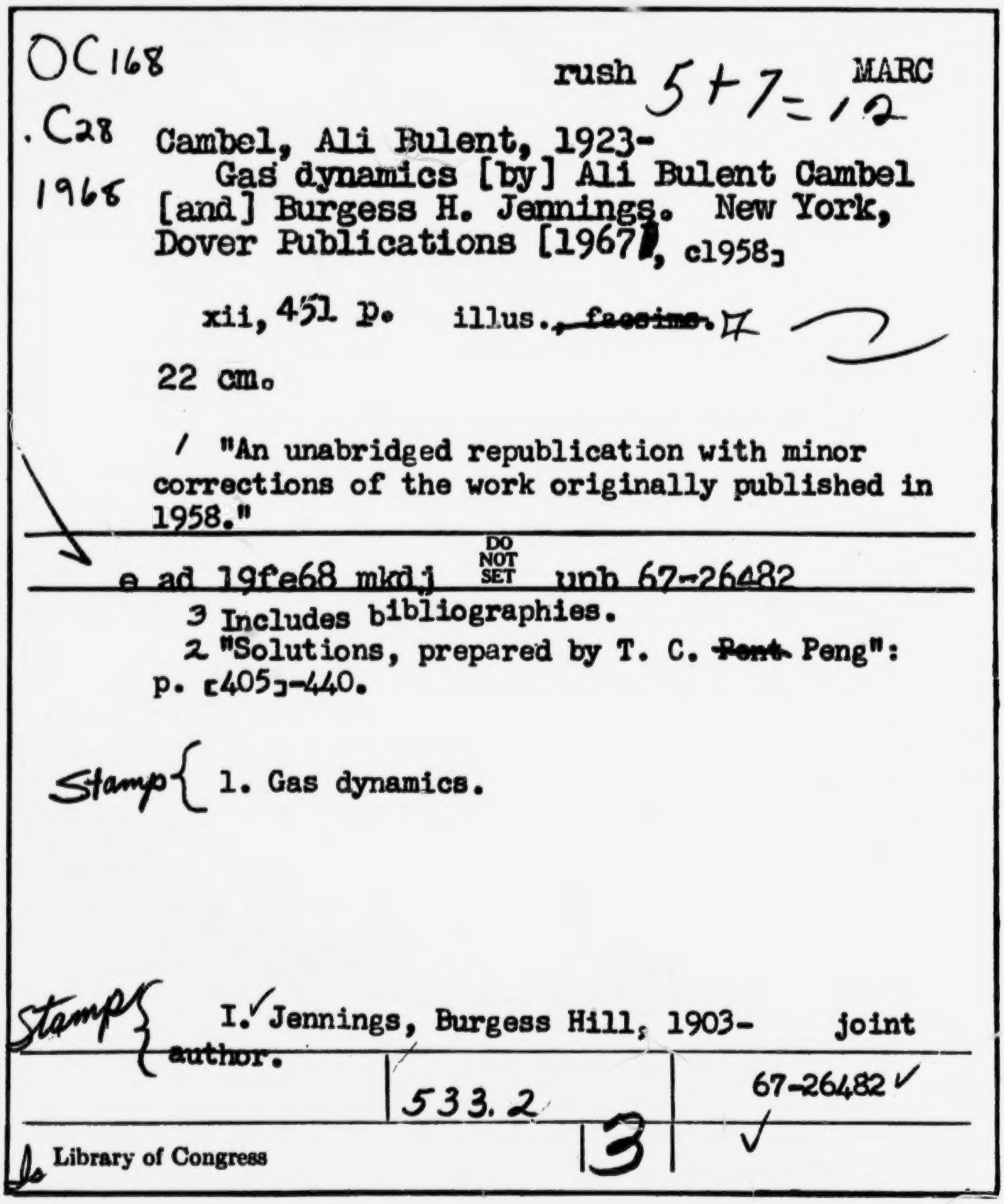
Under the MARC Pilot System and continuing into the MARC Interim System, the manuscript card (Figure 2) prepared by the Library's Processing Department is reproduced on a printed form to serve as the input worksheet for the MARC project (Figure 3). The input worksheet is organized into three parts:

1. The fixed-field data entered into 18 preprinted boxes by the MARC editor.
2. The variable field data consisting of the information on the original manuscript card,
3. The control boxes containing initials and dates filled in by the editors and paper tape punchers as each step in the process is completed.

During the course of the project, the original worksheet evidenced certain inefficiencies that were corrected by redesigning it. In the initial worksheet, the preprinted boxes for the fixed fields were placed at the top of the page. Since it was necessary for the editor to scan the variable field information to complete the fixed fields, it was found more efficient to place these boxes at the bottom of the worksheet where they could be filled in after the variable fields were tagged. The variable field tags used for a guide and a check to the editor were originally placed in the right-hand margin. Since most of the editors were right-handed, this area was covered by the editor's wrist and arm. The new design placed the tag list in the left margin (Figure 4).

Evaluation of the use of a reproduced version of the manuscript card, with the additional data required for the machine-readable record added by an editor, was not possible because a basic constraint of the system in its original definition was not to interfere with the normal workflow of the LC Processing Department. It was always assumed that if the pilot was a success, any operational system would include the redesign of the collection procedures. Under the MARC System the aim is to have the cataloger identify those data elements that must be supplied by a professional librarian or that can only be supplied with the book in hand. The MARC II format has added to the record a number of new elements based on the premise that the MARC process will begin with the cataloger. Future expansion to foreign language material will also necessitate a cataloger with competence in a given language.

One element of the new collection procedures will be a new input worksheet. This worksheet will again be based on the manuscript card but with a number of modifications. There are



OC168 rush 5+7=12 MARC
.C28 Cambel, Ali Bulent, 1923-
1968 Gas dynamics [by] Ali Bulent Cambel
[and] Burgess H. Jennings. New York,
Dover Publications [1967, c1958]

xii, 451 p. illus.~~, facsim.~~

22 cm.

/ "An unabridged republication with minor corrections of the work originally published in 1958."

e ad 19fe68 mkdj DO NOT SET unb 67-26482

3 Includes bibliographies.
2 "Solutions, prepared by T. C. ~~Pent.~~ Peng": p. [405]-440.

Stamp { 1. Gas dynamics.

Stamp { I. ✓Jennings, Burgess Hill, 1903- joint author.

533.2 | 3 | 67-26482 ✓ ✓

Library of Congress

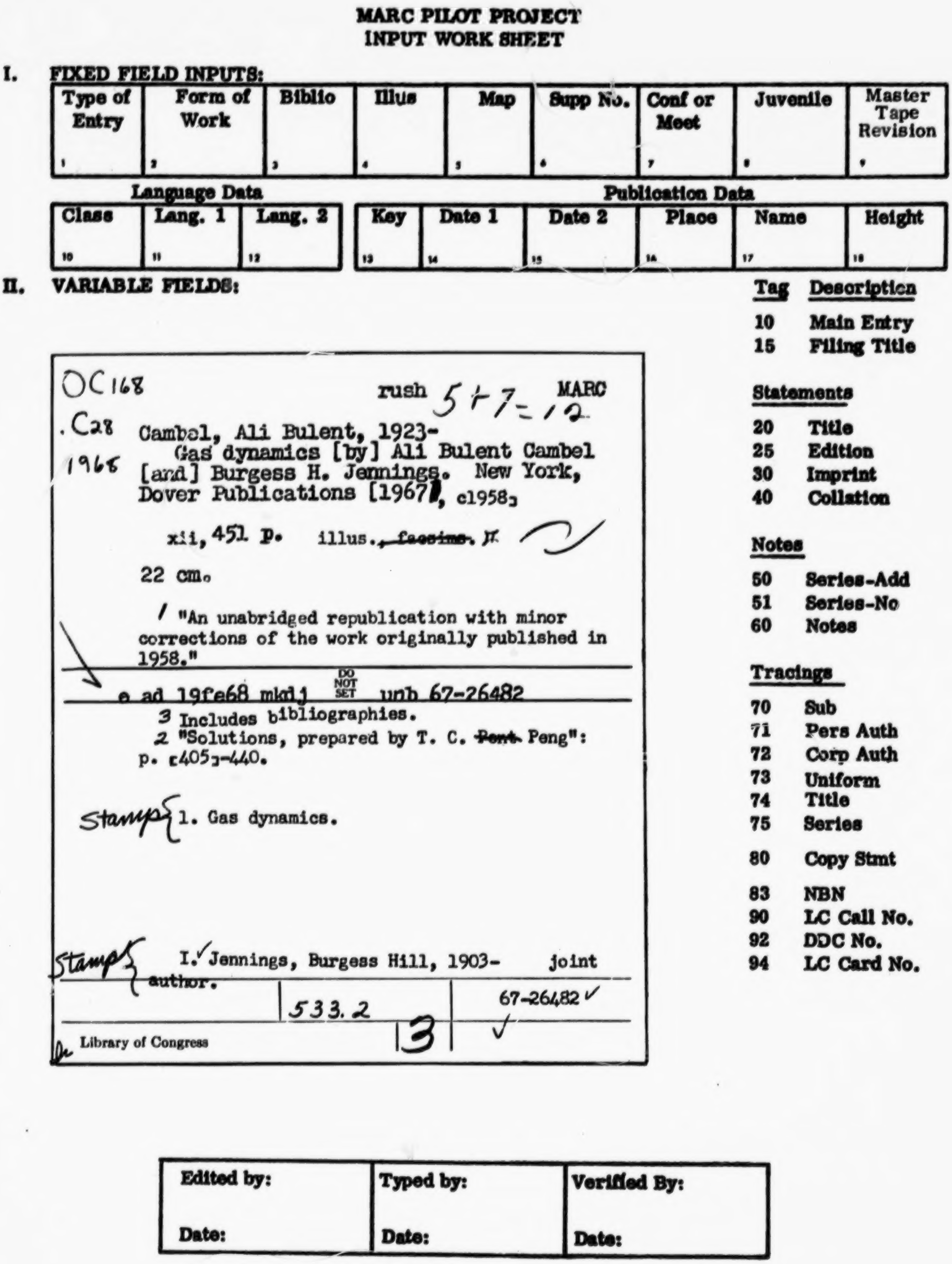
FIGURE 2.—*Manuscript Card*

several stringent requirements placed upon the design of forms by the present methodology of cataloging, processing, and printing in the Library of Congress. The criteria that must be met are as follows:

1. The method used to keyboard and produce the worksheet must also produce, as a byproduct, 3x5 preliminary cards used for the Process Information File and as new acquisitions notices for other parts of the Library. This means that the original preliminary record—including main entry, title, imprint, pagination and series—must be formatted to fit on a 3x5 card. In addition, the worksheet must be of such dimensions as to fit on the copy flow equipment presently used to produce both the 3x5 cards and the original manuscript card.

2. The size of that part of the new worksheet that will be used by the Government Printing Office (GPO) for typesetting cannot greatly exceed the 6x5 size of the current manuscript card due to limitations in the GPO equipment.

3. Additional marks used for machine purposes, i.e., editor's symbols, tags for identifying data elements, etc., must be kept to an



MARC PILOT PROJECT
INPUT WORK SHEET

I. **FIXED FIELD INPUTS:**

Type of Entry	Form of Work	Biblio	Illus	Map	Supp No.	Conf or Meet	Juvenile	Master Tape Revision
1	2	3	4	5	6	7	8	9

Language Data			Publication Data					
Class	Lang. 1	Lang. 2	Key	Date 1	Date 2	Place	Name	Height
10	11	12	13	14	15	16	17	18

II. **VARIABLE FIELDS:**

OC168
.C28
1968

rush 5+7=12 MARC

Cambel, Ali Bulent, 1923-
Gas dynamics [by] Ali Bulent Cambel [and] Burgess H. Jennings. New York, Dover Publications [1967], c1958]

xii, 451 p. illus.~~, facsims.~~

22 cm.

/ "An unabridged republication with minor corrections of the work originally published in 1958."

e ad 19fe68 mkdj DO NOT SET unb 67-26482

3 Includes bibliographies.
2 "Solutions, prepared by T. C. ~~Pent~~ Peng": p. [405]-440.

Stamp { 1. Gas dynamics.

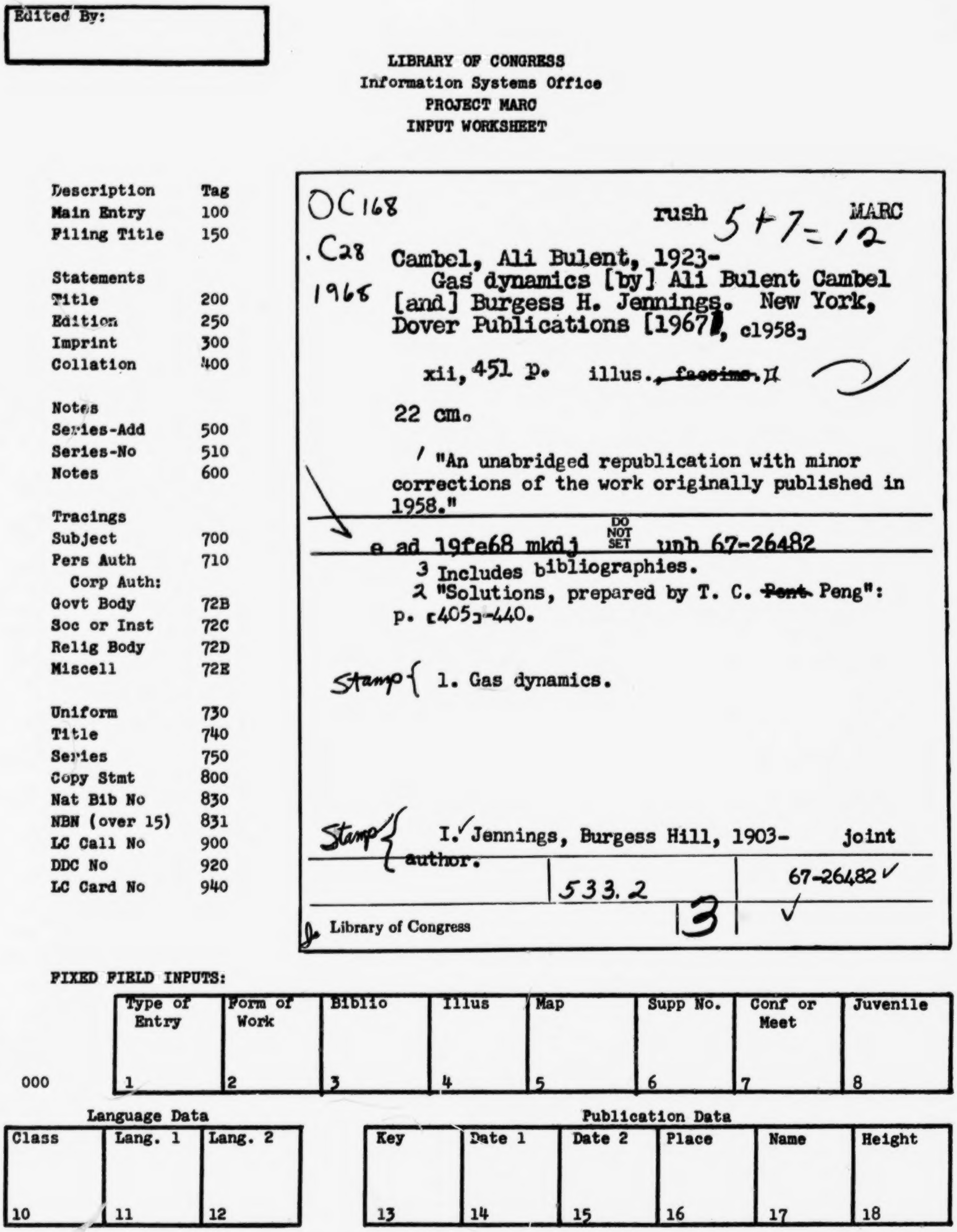
Stamp { I.✓ Jennings, Burgess Hill, 1903- joint author.

	533.2	67-26482 ✓
Library of Congress	3	✓

Tag	Description
10	Main Entry
15	Filing Title
Statements	
20	Title
25	Edition
30	Imprint
40	Collation
Notes	
50	Series-Add
51	Series-No
60	Notes
Tracings	
70	Sub
71	Pers Auth
72	Corp Auth
73	Uniform
74	Title
75	Series
80	Copy Stmt
83	NBN
90	LC Call No.
92	DDC No.
94	LC Card No.

Edited by:	Typed by:	Verified By:
Date:	Date:	Date:

FIGURE 3.—*Original Input Worksheet*



Edited By:

LIBRARY OF CONGRESS
Information Systems Office
PROJECT MARC
INPUT WORKSHEET

Description	Tag
Main Entry	100
Filing Title	150
Statements	
Title	200
Edition	250
Imprint	300
Collation	400
Notes	
Series-Add	500
Series-No	510
Notes	600
Tracings	
Subject	700
Pers Auth	710
Corp Auth:	
Govt Body	72B
Soc or Inst	72C
Relig Body	72D
Miscell	72E
Uniform	730
Title	740
Series	750
Copy Stmt	800
Nat Bib No	830
NBN (over 15)	831
LC Call No	900
DDC No	920
LC Card No	940

QC168 rush 5+7=12 MARC

.C28 Cambel, Ali Bulent, 1923-
1968 Gas dynamics [by] Ali Bulent Cambel [and] Burgess H. Jennings. New York, Dover Publications [1967], c1958]

xii, 451 p. illus., ~~facsims.~~]

22 cm.

1 "An unabridged republication with minor corrections of the work originally published in 1958."

~~e ad 19fe68 mkdj~~ DO NOT SET ~~unh 67-26482~~

3 Includes bibliographies.
2 "Solutions, prepared by T. C. ~~Pont~~ Peng": p. [405]-440.

Stamp{ 1. Gas dynamics.

Stamp{ I. ✓Jennings, Burgess Hill, 1903- joint ~~author.~~

533.2 | 67-26482 ✓

Library of Congress 13 ✓

FIXED FIELD INPUTS:

000

Type of Entry	Form of Work	Biblio	Illus	Map	Supp No.	Conf or Meet	Juvenile
1	2	3	4	5	6	7	8

Language Data

Class	Lang. 1	Lang. 2
10	11	12

Publication Data

Key	Date 1	Date 2	Place	Name	Height
13	14	15	16	17	18

11-19 (rev 7/67)

FIGURE 4.—*Revised Input Worksheet*

absolute minimum on the manuscript card used as copy by the Government Printing Office.

4. Whatever method of collection is implemented must be applicable to the procedures used in the Library's overseas offices and in the Shared Cataloging Division.

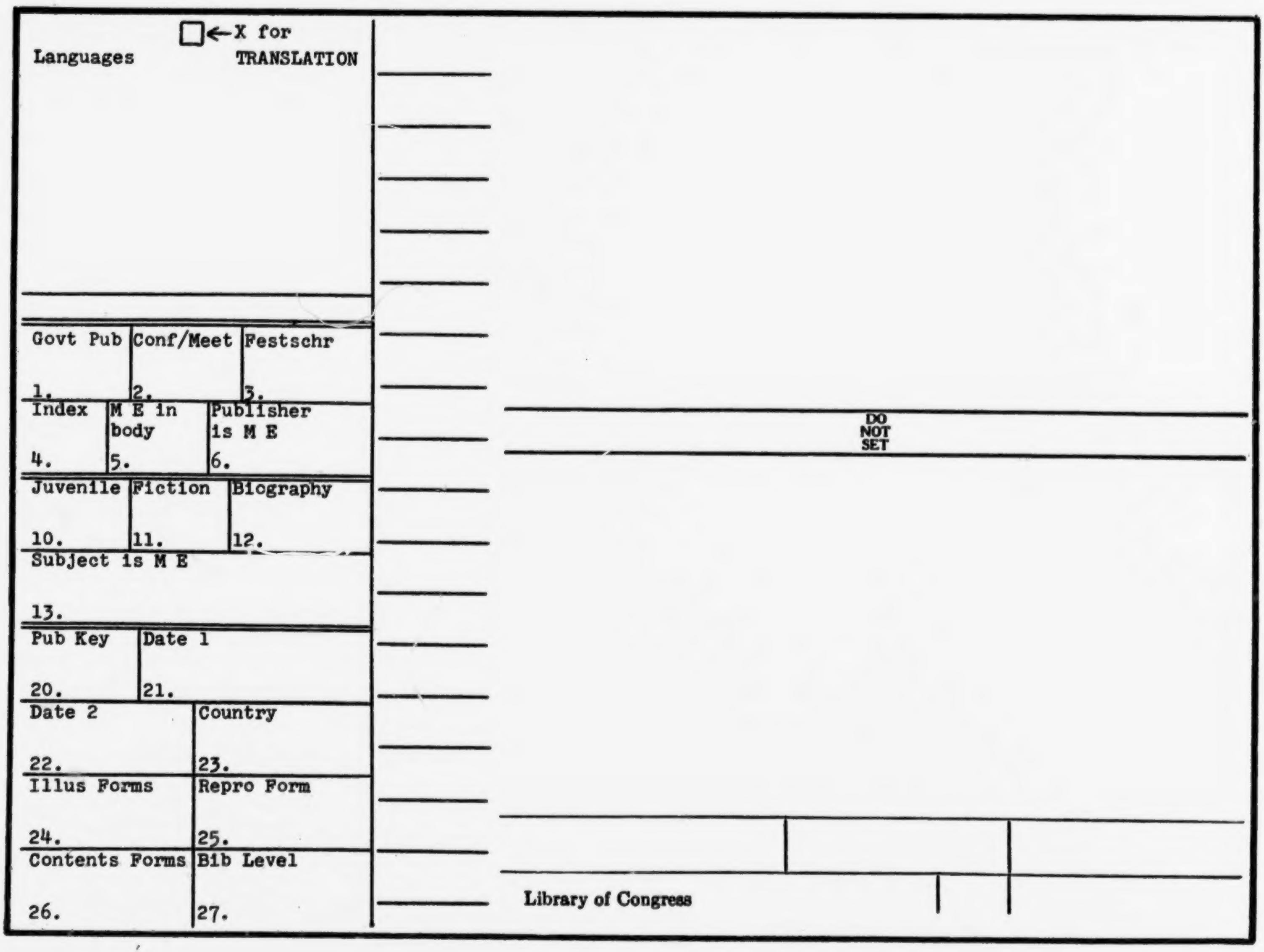
In addition to the above, the form must satisfy all the requirements of tagging and delimiting for the transcription of data in machine-readable form.

Several worksheets have been designed, tested, and eliminated as not meeting the criteria of one or more of the steps in the overall process. Up to the time this is written, all attempts to bypass the reproduction of the original worksheet for the MARC source data before it is forwarded to GPO for printing have failed. Efforts are continuing in this area. However, to meet the schedule for the MARC Distribution Service, an input worksheet has been designed that does meet the criteria 1–4 above. This worksheet will be subjected to an extensive test by the Processing Department. The worksheet (Figure 5) has a perforated section so that the manuscript card section can be removed after reproduction and forwarded to GPO. The copied worksheet, intact, will be the MARC source data.

If the worksheet proves to be effective, it will be used for the cataloging process, and reproduced for the MARC System.

CROSS-REFERENCE TRACING RECORDS

Cross-reference tracings are recorded in the LC processing system on authority cards. This information was supplied to the MARC staff by the cataloging divisions, either by way of re-



Languages ☐←X for TRANSLATION

Govt Pub 1.	Conf/Meet 2.	Festschr 3.
Index 4.	M E in body 5.	Publisher is M E 6.
Juvenile 10.	Fiction 11.	Biography 12.
Subject is M E 13.		
Pub Key 20.	Date 1 21.	
Date 2 22.	Country 23.	
Illus Forms 24.	Repro Form 25.	
Contents Forms 26.	Bib Level 27.	

DO NOT SET

Library of Congress

FIGURE 5.—*Proposed Input Worksheet*

produced copies of the authority cards or on 3x5 typed cards. These cards were used by the editors in the MARC Pilot System as input worksheets. The worksheet for processing such records under the MARC System has not yet been determined.

Data Preparation

Data preparation is defined as those procedures necessary to prepare the data to be transcribed into machine-readable form (editing) and to control the workflow of these data records (control). Data preparation is handled by the MARC System Production Group (MSPG) composed of MARC staff responsible for the production of MARC records from receipt of the worksheet to the verification of the record on magnetic tape. This responsibility includes the control of the workflow, the editing of worksheets, the coding of data elements (e.g., place of publication, language, and publisher), the punching of the data into machine-readable form, and the proofing of machine-generated diagnostic listings against the original worksheets. The staff includes librarians, editors, and papertape typists. In addition to performing all duties associated with training and managing a group of people, the librarians collect and analyze data for cost accounting and test all new techniques designed to improve MARC. MSPG is the laboratory for MARC. A similar group will be established within the Library's Processing Department for data preparation for the MARC Distribution Service.

Control

Because of the sizable quantity of bibliographic data being handled in the MARC pilot, it was thought essential that explicit control procedures be used to monitor and integrate the various operations that were performed. A function designated as MARC Control provided a central distribution and collection point for the data being processed. MARC Control encompassed both the personnel and methods used for routing, scheduling, and monitoring of MARC bibliographic information from its receipt in manuscript form to its delivery in final magnetic tape format.

In the course of the project these procedures changed from a rigid system requiring a large amount of manual handling to a relatively informal system in which most of the control was handled by the computer.

MARC PILOT SYSTEM

The original control procedure under the MARC Pilot System was as follows. Each day a group of manuscript cards was reproduced on input worksheets and forwarded to MSGP. The worksheets were logged in, and each LC card number was recorded in a logbook in numerical sequence by the staff member designated as a control clerk. As the worksheets went through the process, some were rejected for poor reproduction, characters the system could not accommodate, questions for which the editor needed to have the book in hand, or similar problems. These rejections were resubmitted to the control clerk and the numbers deleted from the logbook. The purpose of this process was to indicate which records actually became part of the MARC data base. Since the input worksheets did not arrive at MSPG in any order, the sequencing and copying of each LC card number in the log became an overhead burden to the system that could not be justified. The system was actually bogged down in the attempt to identify the inclusion or exclusion of a record in the system. Therefore, the process of logging records by LC card number was eliminated, but the rejected records were filed and analyzed later in planning the MARC Distribution Service.

Under the revised procedures, each day's worksheets were divided into batches of 100. Within each batch, the records were arranged in numerical sequence by LC card number, and these numbers were transcribed on a batch control sheet which accompanied the batch through the editing process. Each batch was assigned a control number that was recorded in a batch control log. This procedure was used for both new MARC records and corrected records. Each batch of 10 was returned to the control clerk after each step, e.g., editing, punching, in the processing cycle was completed. The control clerk recorded the completion of each step and the date in the batch log. This allowed each batch to be located as it moved through the system.

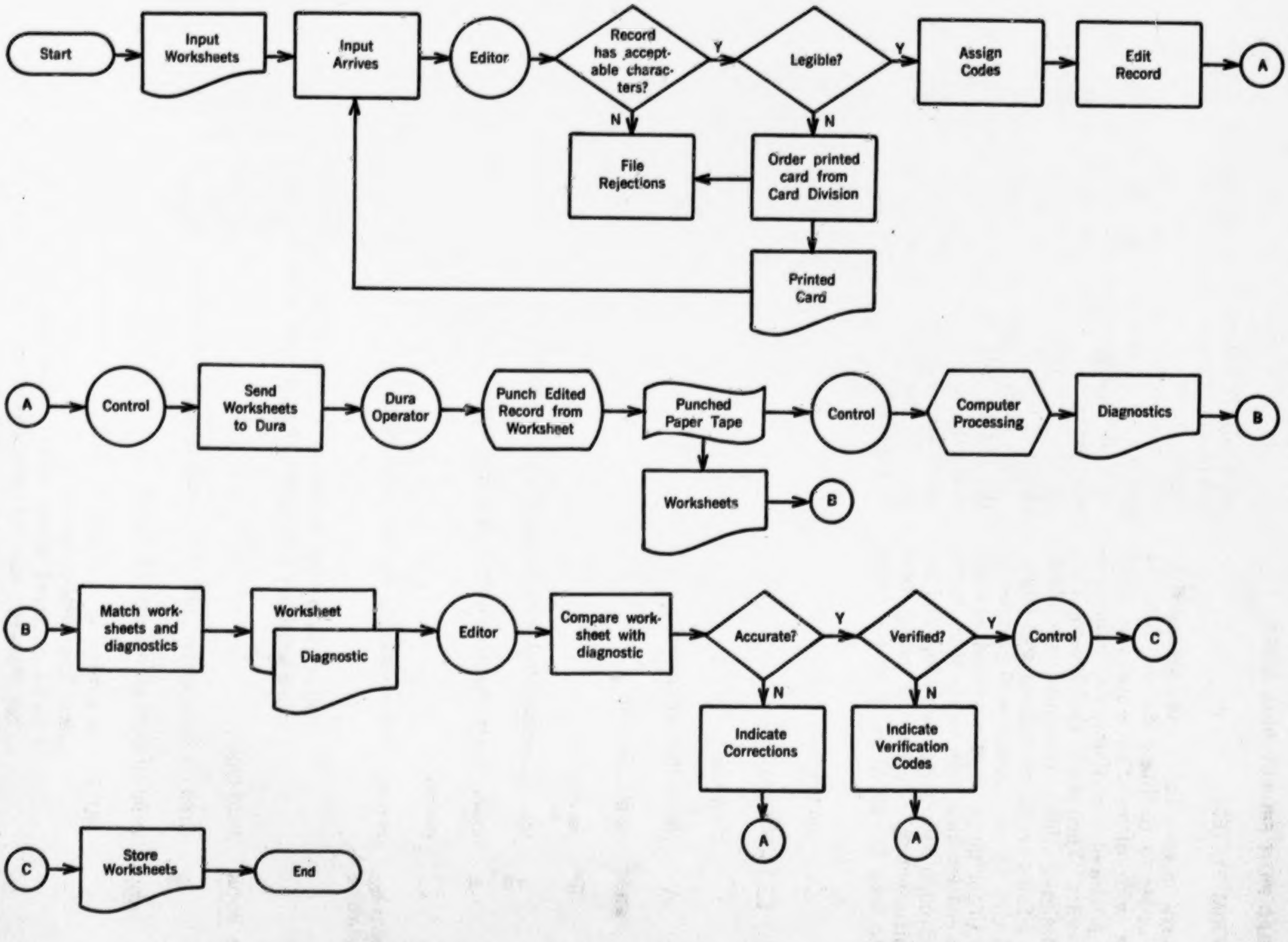
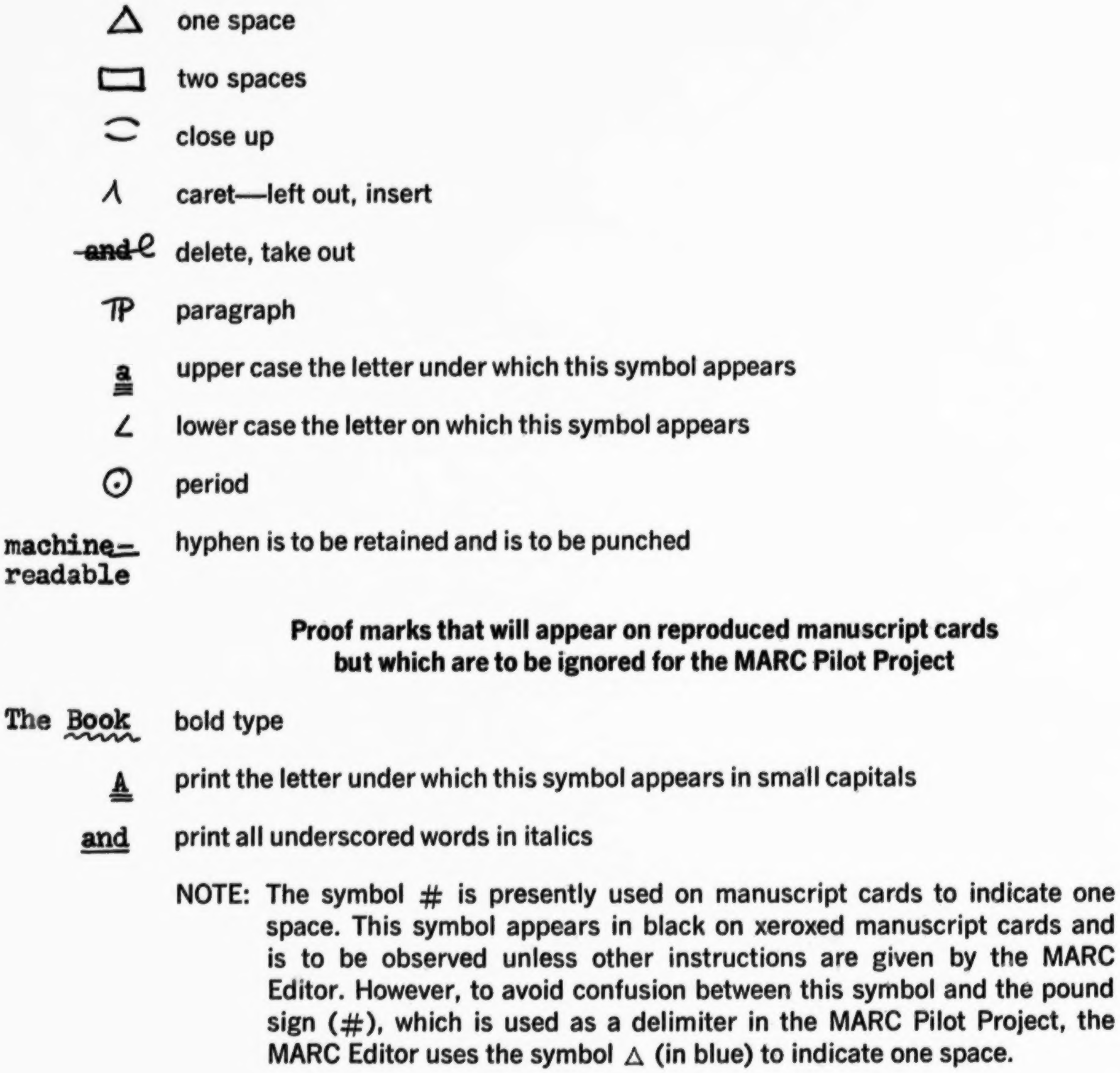

Start
Input Worksheets
Input Arrives
Editor
Record has acceptable characters?
Y
N
Legible?
Y
N
Assign Codes
Edit Record
A
File Rejections
Order printed card from Card Division
Printed Card
A
Control
Send Worksheets to Dura
Dura Operator
Punch Edited Record from Worksheet
Punched Paper Tape
Control
Computer Processing
Diagnostics
B
Worksheets
B
B
Match worksheets and diagnostics
Worksheet
Diagnostic
Editor
Compare worksheet with diagnostic
Accurate?
Y
N
Verified?
Y
N
Control
C
Indicate Corrections
A
Indicate Verification Codes
A
C
Store Worksheets
End

FIGURE 6.—*The MARC Interim System*

MARC INTERIM SYSTEM

The records in each batch were arranged in numerical order to facilitate the matching of worksheets with diagnostics since the print programs produced the diagnostic listings in numerical order. When the MARC Interim System was designed, this requirement was eliminated and a third set of procedures was implemented. MARC records, though still grouped in batches of 10 for purposes of workload, are left in random order within each batch. Each batch is routed through editing and punching, and the only requirement placed on the paper tape typist is to keep the batch of 10 in the order punched. This constraint guarantees that the computer-generated diagnostics, which are output in the order of input, can easily be matched with the worksheets to facilitate the proofing process by the MARC editor. The batch log has been eliminated, reducing the time of the entire processing cycle. Daily logs are now produced by the machine indicating the number and status of each record in process.

The control procedures implemented under the MARC Interim System are basically those that will be used under the MARC System. A schematic diagram of the flow of material through the MARC Interim System is shown in Figure 6.



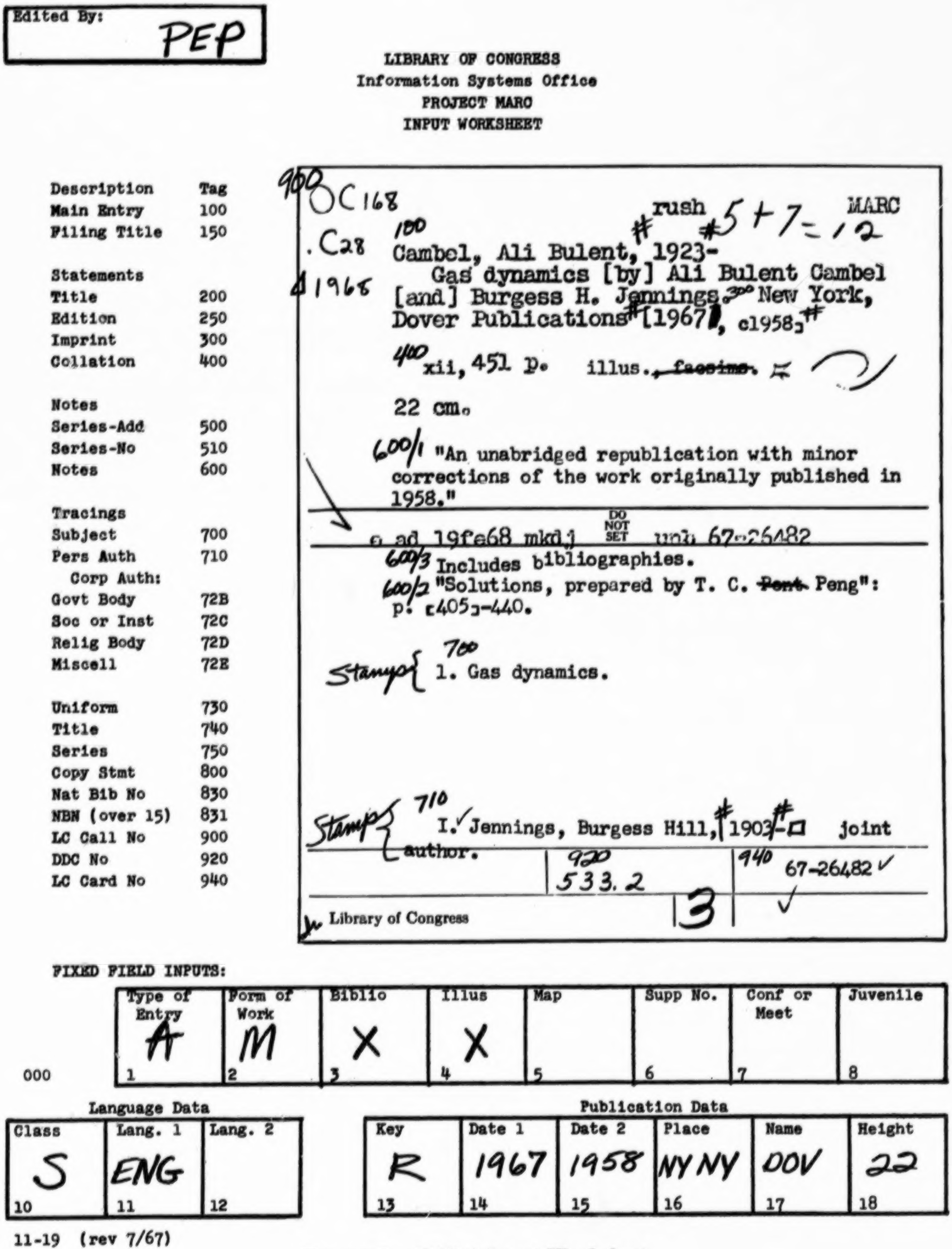
△	one space
▭	two spaces
⁀	close up
∧	caret—left out, insert
~~and~~ ℓ	delete, take out
¶	paragraph
a	upper case the letter under which this symbol appears
∠	lower case the letter on which this symbol appears
⨀	period
machine= readable	hyphen is to be retained and is to be punched

Proof marks that will appear on reproduced manuscript cards but which are to be ignored for the MARC Pilot Project

The Book	bold type
A	print the letter under which this symbol appears in small capitals
and	print all underscored words in italics

NOTE: The symbol # is presently used on manuscript cards to indicate one space. This symbol appears in black on xeroxed manuscript cards and is to be observed unless other instructions are given by the MARC Editor. However, to avoid confusion between this symbol and the pound sign (#), which is used as a delimiter in the MARC Pilot Project, the MARC Editor uses the symbol △ (in blue) to indicate one space.

FIGURE 7.—*Editor's Marks*



Edited By: PEP

LIBRARY OF CONGRESS
Information Systems Office
PROJECT MARC
INPUT WORKSHEET

Description	Tag
Main Entry	100
Filing Title	150
Statements	
Title	200
Edition	250
Imprint	300
Collation	400
Notes	
Series-Add	500
Series-No	510
Notes	600
Tracings	
Subject	700
Pers Auth	710
Corp Auth:	
Govt Body	72B
Soc or Inst	72C
Relig Body	72D
Miscell	72E
Uniform	730
Title	740
Series	750
Copy Stmt	800
Nat Bib No	830
NBN (over 15)	831
LC Call No	900
DDC No	920
LC Card No	940

900 QC168
.C28
1968

rush #5+7=12 MARC

100 Cambel, Ali Bulent,# 1923-
Gas dynamics [by] Ali Bulent Cambel [and] Burgess H. Jennings. 300 New York, Dover Publications# [1967], c1958#

400 xii, 451 p. illus. ~~, facsims.~~

22 cm.

600/1 "An unabridged republication with minor corrections of the work originally published in 1958."

o ad 19fe68 mkdj DO NOT SET unh 67-26482

600/3 Includes bibliographies.
600/2 "Solutions, prepared by T. C. ~~Pent.~~ Peng": p. [405]-440.

Stamps { 700 1. Gas dynamics.

Stamps { 710 I. ✓ Jennings, Burgess Hill,# 1903-# ☐ joint author.

920 533.2 | 3 | 940 67-26482 ✓ ✓

Library of Congress

FIXED FIELD INPUTS:

	Type of Entry	Form of Work	Biblio	Illus	Map	Supp No.	Conf or Meet	Juvenile
000	A	M	X	X				
	1	2	3	4	5	6	7	8

Language Data

Class	Lang. 1	Lang. 2
S	ENG	
10	11	12

Publication Data

Key	Date 1	Date 2	Place	Name	Height
R	1967	1958	NY NY	DOV	22
13	14	15	16	17	18

11-19 (rev 7/67)

FIGURE 8.—*Edited Input Worksheet*

Editing

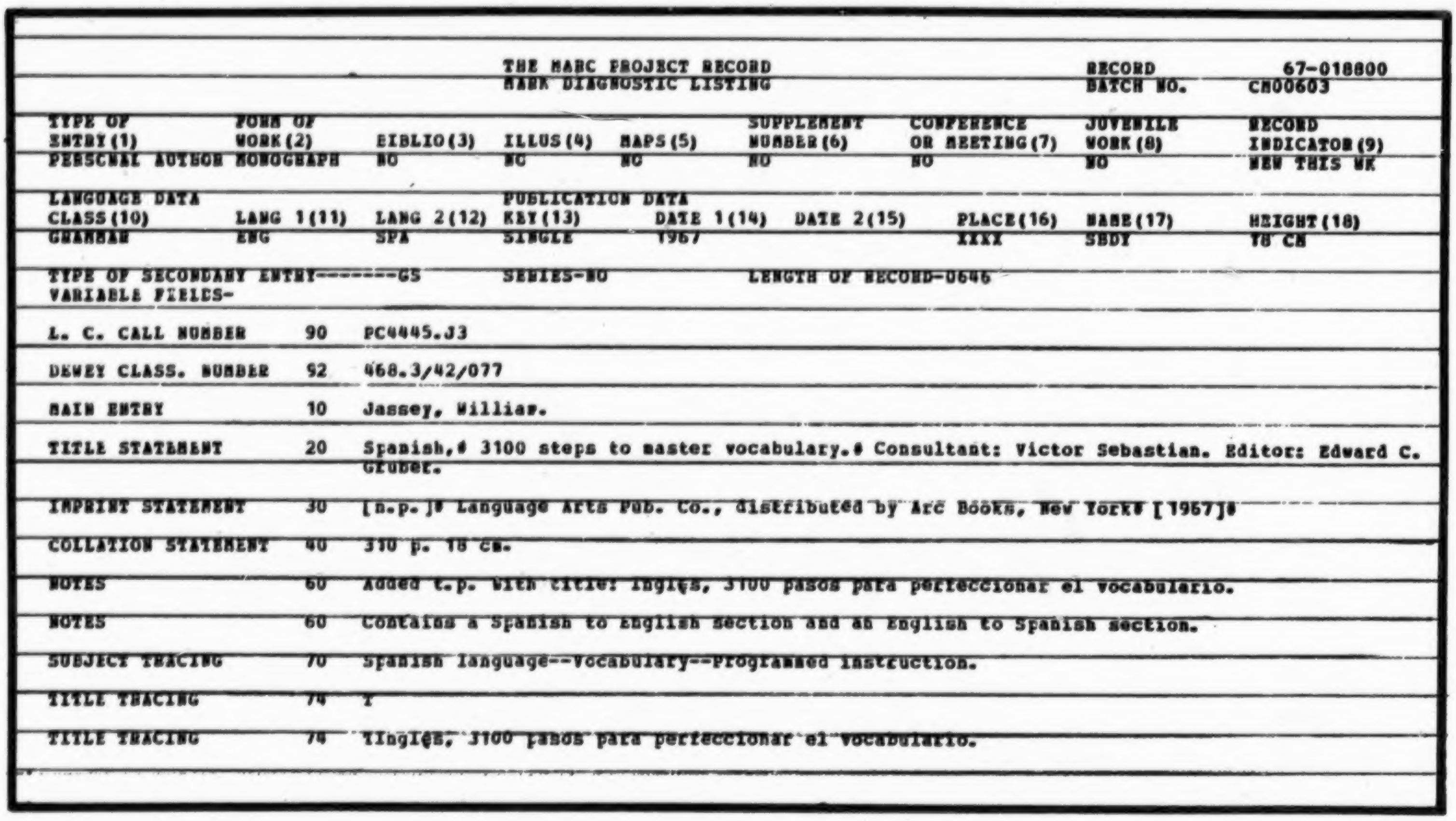
The editing of new input worksheets is the first major step in the conversion of bibliographic data into machine-readable form. The editor performs the following operations:

1. Tags each of the variable fields and inserts delimiters where necessary.
2. Edits and reformats the contents of the variable fields as required. (See Figure 7 for editor's marks.)
3. Enters pertinent fixed field data in the block provided on the form.
4. Proofreads worksheets which have been edited by other editors. (This step was dropped in January 1967 and the proofreading was done when the data on the printed diagnostic listings were verified. It was resumed in July of that year when it became apparent that the final tapes contained too many errors.)
5. After all editing is complete, initials and dates the worksheets and forwards them (in batches of 10) for punching on paper tape. An example of an edited worksheet is shown in Figure 8.

After the data is punched (see section on Data Transcription) the paper tape is input to the computer and transferred to magnetic tape. During this process, a printout diagnostic listing is produced to be used in verifying the data on the tape.

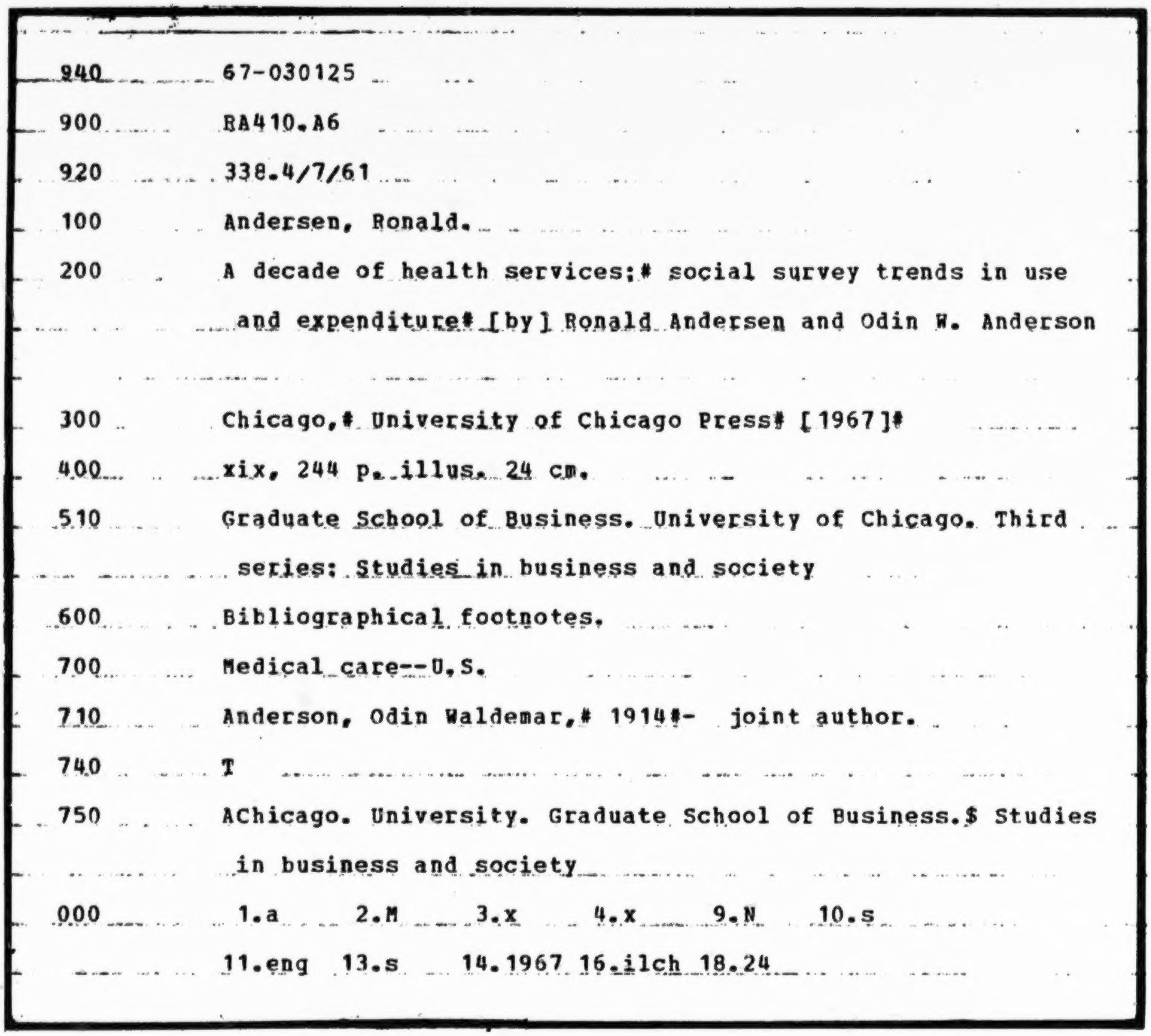
The diagnostics are then matched with the worksheets and compared item by item. In the original pilot system, the diagnostic (see Figure 9) was cumbersome to work with because it did not mirror the worksheet. For example, the presence of a bibliography was indicated on the worksheet by an X in box 3 marked "Biblio," while in the corresponding diagnostic, a "Yes" would appear under the heading "Bibliography." Consequently the diagnostic was redesigned to reflect both the order and the language of the worksheet. The same symbols are used on both input and output, making the diagnostic more useful for proofreading. (See Figure 10.)

In the original MARC Pilot System, the diagnostics were generated within each batch in numerical order. Each diagnostic was then attached to its corresponding worksheet and the two remained together through the remainder of the system. When proofreading the material, the editor removed those worksheets and associated diagnostics needing correction from



```
                                        THE MARC PROJECT RECORD                                   RECORD         67-018800
                                        MARC DIAGNOSTIC LISTING                                   BATCH NO.      CM00603

TYPE OF          FORM OF                                                 SUPPLEMENT    CONFERENCE           JUVENILE    RECORD
ENTRY(1)         WORK(2)         BIBLIO(3)   ILLUS(4)   MAPS(5)          NUMBER(6)     OR MEETING(7)        WORK(8)     INDICATOR(9)
PERSONAL AUTHOR  MONOGRAPH       NO          NO         NO               NO            NO                   NO          NEW THIS WK

LANGUAGE DATA                                PUBLICATION DATA
CLASS(10)        LANG 1(11)      LANG 2(12)  KEY(13)          DATE 1(14)    DATE 2(15)     PLACE(16)   NAME(17)    HEIGHT(18)
GRAMMAR          ENG             SPA         SINGLE           1967                         XXXX        SBDY        18 CM

TYPE OF SECONDARY ENTRY-------GS             SERIES-NO                      LENGTH OF RECORD-0646
VARIABLE FIELDS-

L. C. CALL NUMBER       90    PC4445.J3

DEWEY CLASS. NUMBER     92    468.3/42/077

MAIN ENTRY              10    Jassey, William.

TITLE STATEMENT         20    Spanish,# 3100 steps to master vocabulary.# Consultant: Victor Sebastian. Editor: Edward C.
                              Gruber.

IMPRINT STATEMENT       30    [n.p.]# Language Arts Pub. Co., distributed by Arc Books, New York# [1967]#

COLLATION STATEMENT     40    310 p. 18 cm.

NOTES                   60    Added t.p. with title: Inglés, 3100 pasos para perfeccionar el vocabulario.

NOTES                   60    Contains a Spanish to English section and an English to Spanish section.

SUBJECT TRACING         70    Spanish language--Vocabulary--Programmed instruction.

TITLE TRACING           74    T

TITLE TRACING           74    TInglés, 3100 pasos para perfeccionar el vocabulario.
```

FIGURE 9.—*MARC Project Record Diagnostic Listing*



```
940        67-030125
900        RA410.A6
920        338.4/7/61
100        Andersen, Ronald.
200        A decade of health services;# social survey trends in use
            and expenditure# [by] Ronald Andersen and Odin W. Anderson

300        Chicago,# University of Chicago Press# [1967]#
400        xix, 244 p. illus. 24 cm.
510        Graduate School of Business. University of Chicago. Third
            series: Studies in business and society
600        Bibliographical footnotes.
700        Medical care--U.S.
710        Anderson, Odin Waldemar,# 1914#-  joint author.
740        T
750        AChicago. University. Graduate School of Business.$ Studies
            in business and society
000         1.a    2.M    3.x    4.x    9.N    10.s
           11.eng  13.s   14.1967 16.ilch 18.24
```

FIGURE 10.—*MARC Interim System Diagnostic Listing*

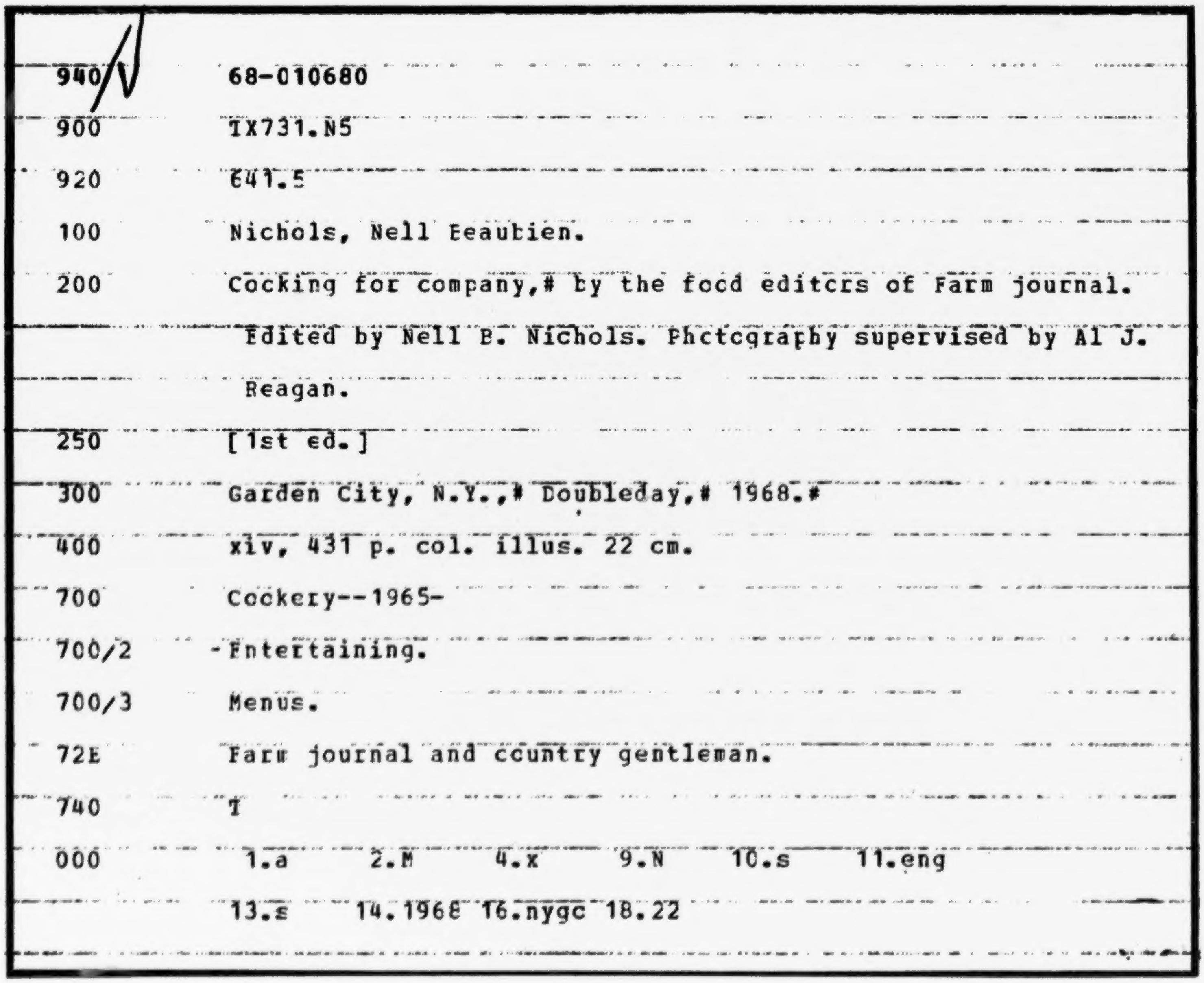
the batch and crossed their LC card numbers off the batch control sheet. The correct or "verified" records remaining in the batch were resubmitted to the tape puncher who punched a list of verified LC card numbers. This tape was input to the computer and the verified records were entered on the MARC master tape.

For the incorrect records, all changes were written out on the diagnostic. These records were then rebatched into "correction batches," and assigned new numbers. The correction batches were returned to the puncher who punched only the corrected fields. When this correction tape was run, the corrections were added to the corresponding unverified records on the magnetic tape, and new diagnostics for each entire record were produced. These were put with their corresponding worksheets and returned to the editor where all procedures were repeated until every record was declared error free.

Under the MARC Interim System control procedures for handling diagnostics have been modified to reduce the problems of batching and rebatching. At the end of a period during which any number of records can be punched, the typist keys 999 into the punched paper tape. The diagnostics are then printed on continuous fanfold rather than one diagnostic per page, and the indication of 999 causes the machine to emit a blank page. The diagnostics are then separated at each blank page, and each set of diagnostics is matched with a group of records. The editor proofs each worksheet with the associated diagnostic and records the status of the record on the diagnostic.

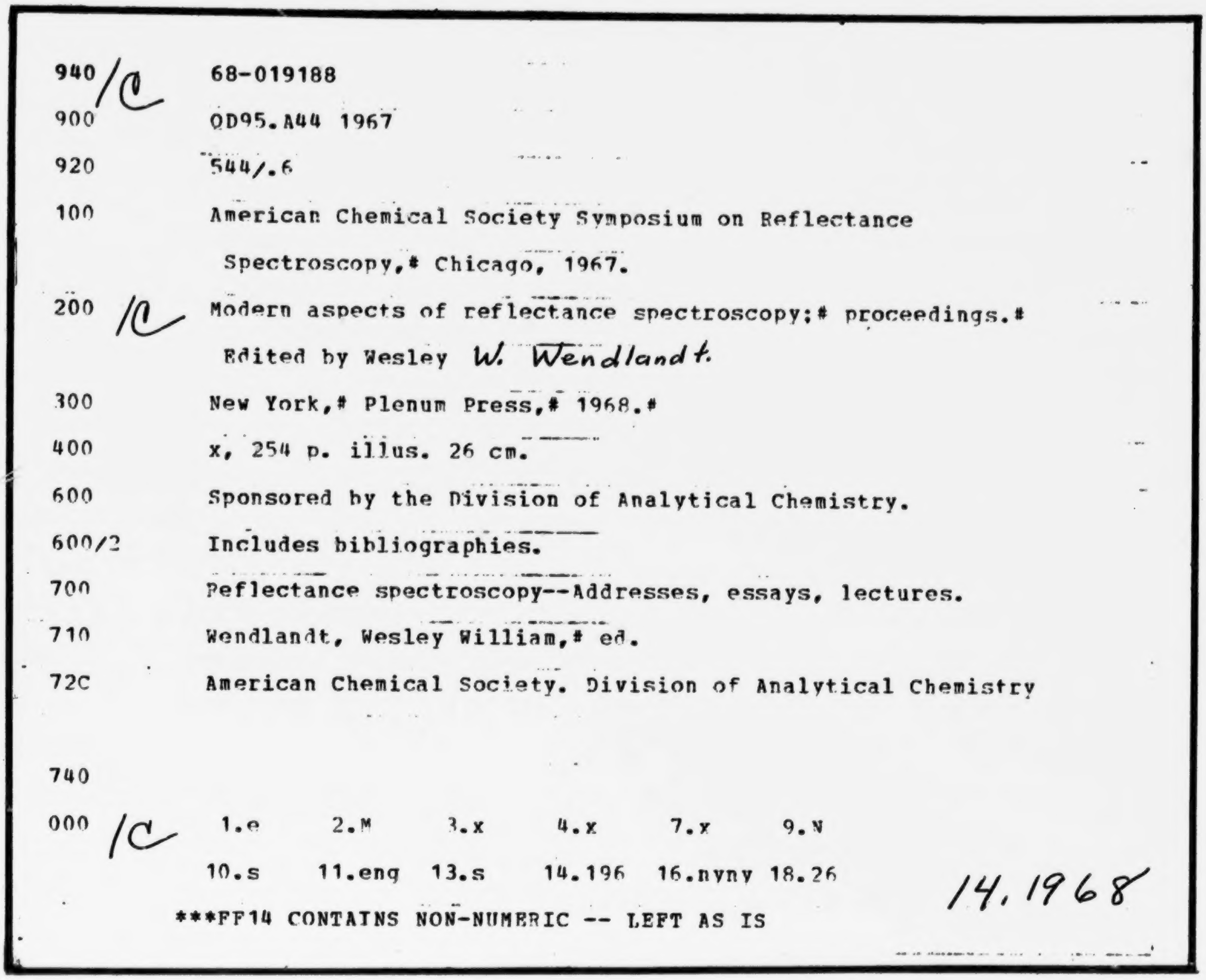
If the record is correct, she records the symbol /V, meaning verified, following the tag for the LC card number. (See Figure 11.) If a record needs a correction, the editor records the symbol /C, meaning correction, tags the field in error, and makes the necessary correction. (See Figure 12.) If the diagnostic is beyond correction (i.e., machine failure, etc.) the editor records the symbol /D meaning delete, (see Figure 13), and the worksheet for that diagnostic is removed and entered into a new batch for repunching.

When this group of diagnostics is completed, it is returned to the puncher. The correction diagnostics produced from this procedure are again printed out on a continuous fanfold. These diagnostics list the verified LC card numbers and the corrected card numbers with those fields which have been changed. The correction diagnostics are compared with the previous diagnostic, and the cycle completed until each record is error free.



```
940/V      68-010680
900        TX731.N5
920        641.5
100        Nichols, Nell Eeaubien.
200        Cocking for company,# by the food editors of Farm journal.
            Edited by Nell B. Nichols. Phctcgraphy supervised by Al J.
            Reagan.
250        [1st ed.]
300        Garden City, N.Y.,# Doubleday,# 1968.#
400        xiv, 431 p. col. illus. 22 cm.
700        Cockery--1965-
700/2     -Entertaining.
700/3      Menus.
72E        Farm journal and ccuntry gentleman.
740        1
000         1.a     2.M     4.x     9.N     10.s     11.eng
            13.s     14.1968 16.nygc 18.22
```

FIGURE 11.—*Verified Diagnostic Listing*



```
940 /C     68-019188
900        QD95.A44 1967
920        544/.6
100        American Chemical Society Symposium on Reflectance
            Spectroscopy,# Chicago, 1967.
200 /C     Modern aspects of reflectance spectroscopy;# proceedings.#
            Edited by Wesley W. Wendlandt.
300        New York,# Plenum Press,# 1968.#
400        x, 254 p. illus. 26 cm.
600        Sponsored by the Division of Analytical Chemistry.
600/2      Includes bibliographies.
700        Reflectance spectroscopy--Addresses, essays, lectures.
710        Wendlandt, Wesley William,# ed.
72C        American Chemical Society. Division of Analytical Chemistry

740
000 /C      1.e      2.M      3.x      4.x      7.x      9.N
           10.s     11.eng   13.s     14.196   16.nyny  18.26      14, 1968
      ***FF14 CONTAINS NON-NUMERIC -- LEFT AS IS
```

FIGURE 12.—*Corrected Diagnostic Listing*

This procedure has proved to be flexible and easy to handle and will be used with slight modifications in the MARC System.

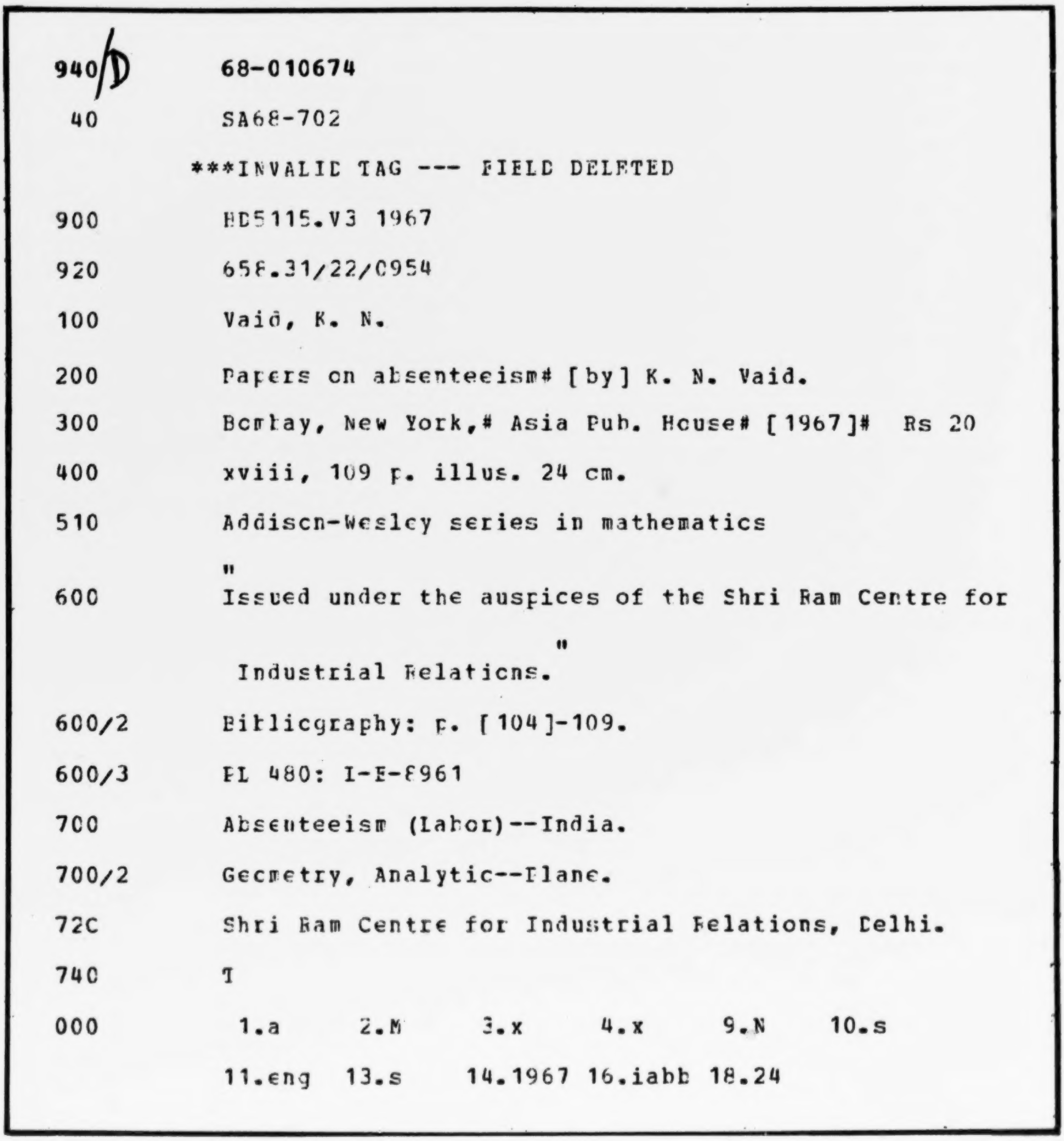
CROSS-REFERENCE TRACING RECORDS

Cross-reference records under the MARC Pilot System were processed in the same manner as bibliographic records. Each record was assigned a control number. After being batched into groups of 10, they were edited and tagged by the MARC editors. They were then punched on paper tape and input to the computer. Diagnostics were produced for verification by the editors. In May 1967, production of these records was suspended pending the development of a more comprehensive cross-reference system.

Data Transcription

Data transcription is the process of putting information into a form that can be read by a machine. The form used in the MARC project was punched paper tape. The decision to use paper tape was based on the following considerations.

1. A survey of libraries in various stages of automation indicated that punched paper tape was more efficient for the processing of variable length bibliographic records than punched cards. The inconvenience of the limitation of



```
940/D       68-010674

 40         SA68-702

         ***INVALID TAG --- FIELD DELETED

900         HD5115.V3 1967

920         658.31/22/0954

100         Vaid, K. N.

200         Papers on absenteeism# [by] K. N. Vaid.

300         Bombay, New York,# Asia Pub. House# [1967]#  Rs 20

400         xviii, 109 p. illus. 24 cm.

510         Addison-Wesley series in mathematics

            "
600         Issued under the auspices of the Shri Ram Centre for
                                          "
             Industrial Relations.

600/2       Bibliography: p. [104]-109.

600/3       PL 480: I-E-8961

700         Absenteeism (Labor)--India.

700/2       Geometry, Analytic--Plane.

720         Shri Ram Centre for Industrial Relations, Delhi.

740         T

000          1.a     2.M     3.x     4.x     9.N     10.s

            11.eng  13.s    14.1967 16.iabb 18.24
```

FIGURE 13.—*Deleted Diagnostic Listing*

the 80-column card and the necessity of linking one card to another by a unique number for field and record seemed to compensate for the added complexity of correction procedures with the use of paper tape.

2. The design of procedures for a paper-tape typewriter could be used with an on-line typewriter without retraining personnel.

3. There were more characters available on a paper-tape typewriter keyboard than on a keypunch keyboard.

The Dura Mach 10 automatic typewriter was chosen as the input device because it utilizes the standard Selectric typing mechanism and satisfied point 2 above.

The Dura Mach 10 typewriter was designed with a standard keyboard* and produced typed copy in addition to punched paper tape. An added feature of the machine was a photoelectric reader which could "read" a program tape and "instruct" the typewriter. Under the control of a prepunched program tape, automatic typing of repetitive information, such as tags and field marks, could be accomplished.

The Dura Mach 10 typewriter used an 8-channel punched paper tape with unique codes for each of the 44 character keys on the typewriter keyboard and for space, backspace, uppercase, lowercase, carriage return, tabulation, error, and tape-stop functions.

Additional control keys, listed below, were used in both manual and automatic typing. The last three codes were designated for specific computer instructions.

Non Print
Print Restore
Tape Skip
Punch Off
Punch On
End of Field (EOF)
End of Record (EOR)
End of Tape (EOT)

MARC PILOT SYSTEM

Procedures for the pilot project utilized the programming capability of the Dura Mach 10. Program tapes were prepared for the transcription of new MARC records, corrections to MARC records, cross-reference tracing records, corrections to cross-reference tracing records, and verified MARC records.

The program tapes were designed to perform three functions: 1) to insert proper codes for the identification of the content of the various fields, 2) to space and position the Dura carriage, and 3) to type the name of each field that could be present in the record as an instruction to the Dura operator. The first program tapes automatically listed all items of fixed field data and variable field data; i.e., if a field was not used, the tag and the name of the field were typed regardless.

As experience was gained, a new program was designed to shorten the typing operation, by making use of the Tape Skip (TS) capability on the typewriter. This feature meant that when the typist punched the Tape Skip key, the program tape was fed through the machine without causing any action until a Tape Skip Restore (TSR) code was reached. Instead of requiring the typing of the name and tag number of each variable field sequentially, this program took advantage of the expected frequency of occurrence of each field in the record to skip over fields used infrequently.

A count was made to determine the occurrence of each variable field in a bibliographic record. The fields fell into three groups: nine of the 20 occurred in nearly every record; three others were frequently used; and the remaining eight occurred infrequently. As the tags for the fields had to be entered in numerical order because of MARC Pilot System computer program specifications, it was possible to use the Tape Skip capability before five frequently used tags to permit a speedup of the typing operation. Table 5 lists each tag together with a schematic indication of the points at which the Tape Skip could be used.

TABLE 5.—*Tape Skip Restore on MARC Program Tapes*

	Tag	*Description*
TS	10	Main Entry
	15	Filing Title
		Statements
TSR	20	Title
TS		
	25	Edition
TSR	30	Imprint
TS		
	40	Collation
		Notes
	50*	Series-Add
	51*	Series-No
TSR	60*	Notes
		Tracings
TS		
	70*	Sub
	71*	Pers Auth
	72*	Corp Auth
	73*	Uniform
TSR	74*	Title
TS		
	75*	Series
	80	Copy Stmt
TSR	83	NBN
	90	LC Call No.
	92	DDC No.
	94	LC Card No.

*Repeatable tags

---

* The design of the MARC keyboard is described in chapter 7.

The program tape routines used were identical to those originally designed, with the addition of the five Tape Skip Restore codes and the corresponding deletion of the name of the variable field. A sample of the hard copy produced on the Dura is shown in Figure 14.

```
00           1.A    2.M       3.      4.    5.      6.       7.        8.        9.

             10.S   11.FRE    12.    13.S   14.1926 15.      16.FRPA   17.XXXX   18.19

10           Laleau, Léon,# 1892#-

20           La flèche au coeur;# poèmes.# Préf. de Maurice Rostand.

30           Paris,# H. Parville# [1926]#

40           122 p. 19 cm.

60

74           T

83

90           PQ2623.A42F5

92

94           50-51832

END OF RECORD.PUSH SKIP IF MORE.
```

FIGURE 14.—*Sample of Dura Hard Copy*

MARC INTERIM SYSTEM

Although in theory program tapes were an aid to the typist, experiences with the MARC Pilot System procedures indicated that in reality, for free form bibliographic data (i.e., data whose length, order, and number of repeatable fields cannot be predetermined), program tapes proved to be a hindrance.

Even though the program tapes were designed to use the Tape Skip feature to bypass fields used infrequently, the fact that several fields for the bibliographic description of monographs (i.e., subject tracings, added entries, etc.) were repeatable, meant that the typist had to enter exception routines for each repeatable tag. A program tape could not be designed to punch an identifying tag automatically when the number of times the tag was needed varied from record to record.

The combination of the TSR and the exception routines resulted in a method of typing MARC records that included too many typewriter strokes used solely for control purposes. In addition, the correction capabilities of the original procedures were very limited, since the backspace was used to signal that the following character was a diacritical. (This technique was used to enable the diacritical to be typed over its associated character on the Dura Mach 10 hard copy.).

As redesigned, the input transcription technique completely eliminates the use of program tapes and has more advanced correction capabilities. The result has been a considerable increase in the number of records produced each day by a typist. The following schematic illustrates the new technique for transcribing new MARC records from the worksheet.[4]

```
C
C
C
ID T   DATA                                 C
ID T   DATA                                    C
T      DATA                              C
T      DATA     C
ID T   DATA  C
ID T   1. DATA T   3. DATA T   2. DATA C
T      7. DATA T   15. DATA C
ID T   DATA C
C
C
C
```

WHERE:

C = A carriage return function code
ID = A set of characters (Tag) used to identify the following unit of information (Data)
T = A tabulation function code
1. = A notation to identify fixed-length fields within a unit of information (Data)
DATA = Any information

AND

Three C's = Beginning or end of record
One C followed by a Non-T = beginning or end of field
One C followed by a T = continuation of field.

The new correction procedures allow the use of the backspace code to delete individual characters, the delete code to delete words, and the delete/carriage return codes to delete an entire field.

EXAMPLES:
ID T The CanterbiBury
Where B = Backspace code
ID T The Canterbiry Tales DD Canterbury Tales.
Where D = delete/per word
ID T The Canterbiry Tales, D C
Where D C = delete field

An example of the hard copy produced in the MARC Interim System is shown in Figure 15.

The revised transcription methods show the following advantages over the original procedures:

1. Batching of records is no longer necessary.
2. Special tape handling to change program tapes for different types of records is no longer necessary.
3. Only normal typing strokes are needed for control purposes, e.g., carriage return, tabulation, backspace, delete.

This method of transcription has proved satisfactory and will serve as the basis of the transcription procedure used in the MARC System.

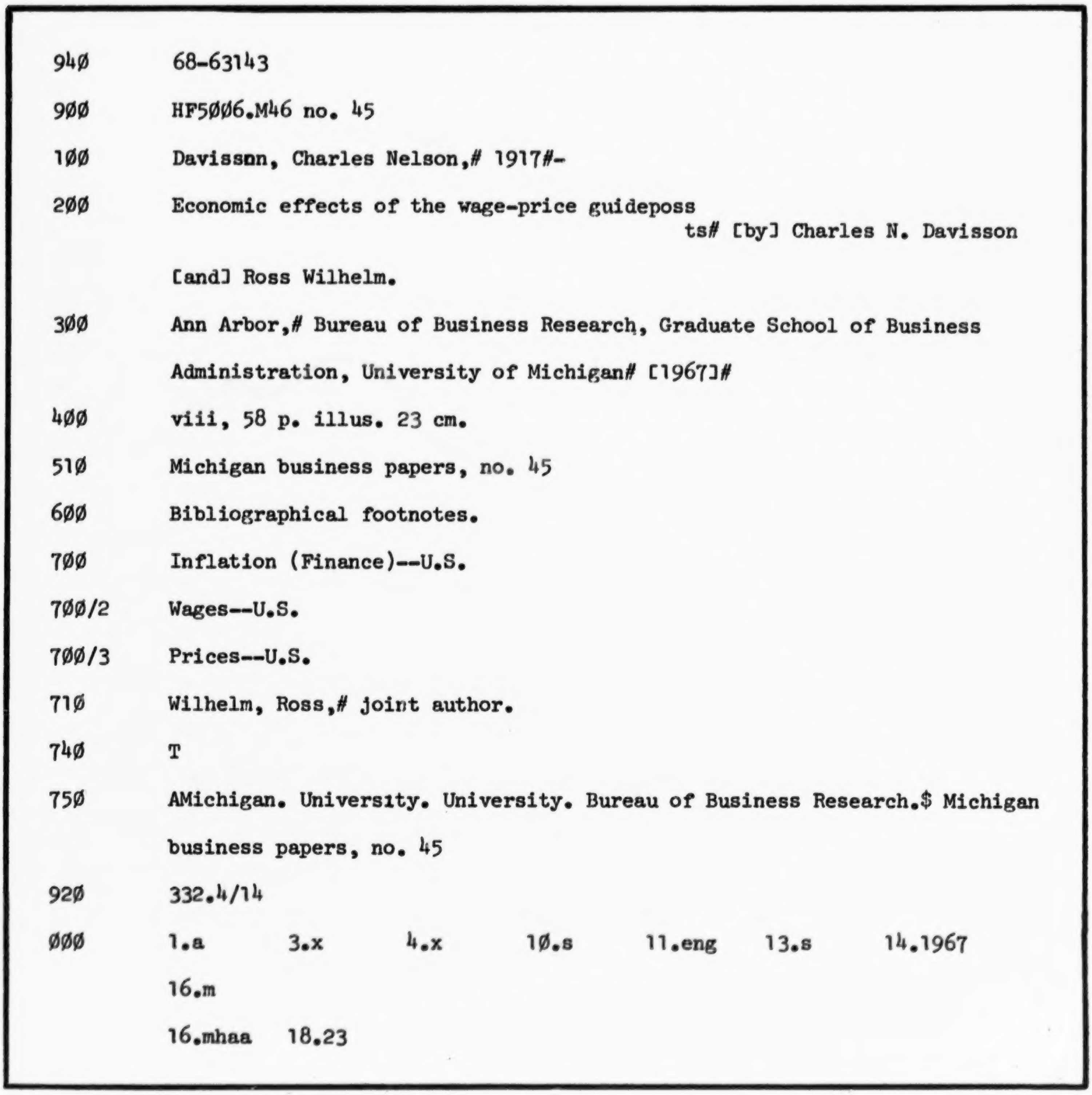
Computer Processing

This section describes the computer processing system in terms of the functions required to accomplish the defined objectives. The overall system comprises a series of tasks to be performed, each of these in turn consisting of a single computer program or a group of programs which collectively perform a specific function.

MARC PILOT SYSTEM

The MARC Pilot Project computer processing system was designed for and implemented on the IBM System/360 Model 30 at the Library of Congress. The programs were written in System/360 assembler language for running under the System/360 Tape Operating System. The programs were grouped into two series of operations, the daily processing cycle and the weekly processing cycle. The daily cycle input, checked, and corrected bibliographic information and converted it to verified records in digital form on magnetic tape. The weekly cycle added these reports to and updated the master file.

The following subsections define the programming system in terms of the functional



```
940       68-63143

900       HF5006.M46 no. 45

100       Davissnn, Charles Nelson,# 1917#-

200       Economic effects of the wage-price guideposs
                                         ts# [by] Charles N. Davisson

          [and] Ross Wilhelm.

300       Ann Arbor,# Bureau of Business Research, Graduate School of Business

          Administration, University of Michigan# [1967]#

400       viii, 58 p. illus. 23 cm.

510       Michigan business papers, no. 45

600       Bibliographical footnotes.

700       Inflation (Finance)--U.S.

700/2     Wages--U.S.

700/3     Prices--U.S.

710       Wilhelm, Ross,# joint author.

740       T

750       AMichigan. University. University. Bureau of Business Research.$ Michigan

          business papers, no. 45

920       332.4/14

000       1.a      3.x      4.x      10.s     11.eng   13.s     14.1967

          16.m

          16.mhaa   18.23
```

FIGURE 15.—*Sample of Dura Hard Copy for MARC Interim System*

operations performed during each processing cycle and describe the flow of information through the individual component programs.

A. Daily Processing Cycle

The Daily Processing Cycle consisted of 10 primary functions. Briefly, Functions 1, 2, and 3 read the punched paper tape records and merged them with data records previously entered; Functions 4 through 9 performed the verification of MARC records and cross-reference tracing records; and Function 10 prepared a diagnostic listing of new or corrected data records to aid the MARC editor in the verification and correction procedures. In the following paragraphs these functions are outlined in terms of individual program applications and record processing flow through the Daily Processing Cycle.

Function 1: TRANSLATE

1. Read paper tape records consisting of new MARC records and corrections to previously processed MARC records, new cross-reference tracing records and corrections to previously processed cross-reference tracing records.

2. Converted Dura-coded input data to Extended Binary Coded Decimal Interchange Code (EBCDIC) for computer processing.

3. Called for appropriate subroutines (depending upon type of data) to format the records.

a. New MARC records—Subroutines FORMAT MARC FIXED FIELD and FORMAT MARC VARIABLE FIELD performed the formatting of the two parts of the MARC record. In addition, the validity of the input data was tested. Errors identified by an error description number were printed.

b. MARC correction records—Subroutine FORMAT CORRECTION MARC RECORD formatted a correction record and performed a limited test of the validity of the input data and identified errors.

c. New cross-reference tracing records—Subroutine FORMAT NEW CROSS-REFERENCE RECORD formatted a new cross-reference record and performed a limited test of the validity of the input data and identified errors.

d. Cross-reference correction records—Subroutine FORMAT CORRECTION CROSS-REFERENCE RECORD formatted a correction cross-reference record and performed a limited test of the validity of the input data and identified errors.

4. As the processing of each record through the appropriate subroutine was completed, the formatted record was written on the MARC Data Input Tape.

Function 2: SORT INPUT DATA

When all of the punched paper tape data had been written on the MARC Data Input Tape, the data were sorted and merged by the SORT/MERGE program. The new MARC records and corrections to previously processed MARC records were arranged by LC card number; the new cross-reference tracing records and corrections to previously processed cross-reference tracing records were arranged by the assigned control number. The MARC records preceded cross-reference tracing records on the output tape (MARC Work Tape).

Function 3: BUILD

1. New records from the sorted MARC Work Tape were merged with the MARC and cross reference-tracing records from the previous Daily Processing Cycle or the Weekly Processing Cycle, respectively, and a new MARC Work Tape was built. During the merge process, correction records from the sorted MARC Work Tape were applied to the pertinent MARC records and cross-reference tracing records.

2. A status indicator was set for the records to which corrections had been applied and for new records. This indicator resulted in a diagnostic printout of these records for proofreading purposes.

Function 4: READ VERIFIED LC NUMBERS

Paper tape records containing LC catalog card numbers of MARC records that had been verified and accepted as correct by the MARC editor were read and the numbers were reformatted. This program also tested the validity of the input data and identified detectable errors.

Function 5: SORT VERIFIED NUMBERS

The verified record numbers were sorted by the SORT/MERGE program into a sequence comparable to that of the MARC records on the MARC Work Tape.

Function 6: UPDATE MARC RECORDS

The MARC Work Tape was searched for those records that were verified. A status-indicator was set for these records resulting in their inclusion as verified records on the updated MARC Work Tape for ultimate input to the Weekly Processing Cycle. The following lists were printed:

1. A list of the LC catalog card number of all the verified MARC records on the updated MARC Work Tape.

2. A list of the LC catalog card numbers for which no match could be found on the MARC Work Tape.

3. A list of LC catalog card numbers of all MARC records on the updated MARC Work Tape that were still unverified.

Function 7: READ VERIFIED CROSS-REFERENCE NUMBERS

Function 8: SORT VERIFIED CROSS-REFERENCE NUMBERS

Function 9: UPDATE CROSS-REFERENCE TRACING RECORDS

Functions 7, 8, and 9 processed cross-reference tracing records in the same manner as functions 4, 5, and 6 for the processing of MARC records.

Function 10: DIAGNOSTIC

The PRINT DIAGNOSTIC program read records from the updated MARC Work Tape and printed a diagnostic listing of the entire content of each MARC record and cross-reference tracing record that was newly entered or corrected by the current day's processing. This program also built the MARC Work Tape which contained MARC records and cross-reference tracing records that had been verified and those that were still being processed (and hence had not yet been accepted as correct). This new MARC Work Tape was used to input to the next Daily Processing Cycle or the Weekly Processing Cycle.

Two status indicators were set in each MARC record and cross-reference tracing record as it moved through the daily and weekly processing functions. The Daily Status Code (DSC) was used in both cycles, whereas the Weekly Status Code (WSC) was applicable to the Weekly Processing Cycle only. The DSC comprised three symbols: P meaning record entered for first time, L meaning diagnostic printed but record not verified, and G meaning verified record. The WSC which was set initially during the daily processing consisted of two symbols: N indicating a new (unique) record and R indicating a revision to a verified record. (Two other WSC were set in the weekly processing indicating the age of the records on the tape: O indicating that the record was one week old and blank indicating that it was two weeks old). In the Daily Processing Cycle, the DSC was used to reflect the status of a record based on three levels of record validation.

Table 6 summarizes the action taken in the daily processing functions based on the Daily Status Code.

TABLE 6.—*Action Taken in Response to Daily Status Code*

	P—Records Entered For First Time	L—Diagnostic Printout But Record Not Verified	G—Verified Record
Function 1 TRANSLATE	All new input records given status code P.	N/A	N/A
Function 3 BUILD	Noncorrected records retain status code P. Correction records applied to corresponding L records.	Noncorrected records retain status code L. Status code of corrected records changes from L to P.	Status code G retained.
Functions 6 and 9 UPDATE MARC RECORDS and UPDATE X-REF RECORDS	Status code P retained.	Status code of newly verified records changes from L to G. Unverified records retain status code L.	Status code G retained.
Function 10 DIAGNOSTIC	Status code changed from P to L	Status code L retained.	Status code G retained.

B. Weekly Processing Cycle

Once each week a series of computer programs updated the master files of information and produced copies of these bibliographic records for the MARC participating libraries. The Weekly Processing Cycle consisted of eight primary functions. Briefly, Function 11 selected the verified data and added them to the master file; Function 12 generated the author/title records; Functions 13, 14, and 15 prepared the data for merging; Functions 16 and 17 merged new data into the master file; and Function 18 duplicated* copies of the updated master files for the participants. In the following paragraphs these functions are outlined in terms of individual program applications and record processing flow through the Weekly Processing Cycle.

---

* Argonne National Laboratory assumed the responsibility for this function when it became evident that this operation was taking too much time because of the particular computer configuration at the Library of Congress.

Function 11a: SEPARATE INPUT RECORDS and Function 11b: LIST VERIFIED NUMBERS

1. The new MARC records whose status indicator had been set at G (verified) were selected and recorded on the New MARC Tape.

2. The new MARC records and cross-reference tracing records that had a status indicator of L (indicating data still being processed and not yet accepted as correct) were selected and recorded on the MARC Work Tape (Residual). This tape was subsequently used as an input to the Daily Processing Cycle. The LC catalog card number of MARC records and the control numbers of cross-reference tracing records written on this tape were printed.

3. The cross-reference tracing records whose status indicators had been set at G were selected and separated into two types, subject and descriptive. Both types of cross-reference tracing records were written as separate files on the Cross-Reference Tape.

4. The LC catalog card numbers of all newly verified MARC records were printed. A diagnostic was also printed for error records.

Function 12: EXTRACT AUTHOR/TITLE DATA

The EXTRACT A/T program read the author or title fields or both from each record on the New MARC Tape, formatted them into a fixed-length record of 134 characters, and recorded these A/T records into one file on the Author/Title Tape. A diagnostic was printed for those records in error.

Function 13: SORT A/T RECORDS

The Author/Title Tape was sorted alphabetically by the SORT/MERGE program and output on the sorted Author/Title Tape.

Function 14: SORT DESCRIPTIVE CROSS-REFERENCE RECORDS

The SORT/MERGE program arranged the descriptive cross-reference tracing records alphabetically on the sorted Descriptive Cross-Reference Tape.

Function 15: SORT SUBJECT CROSS-REFERENCE RECORDS

The SORT/MERGE arranged the subject cross-reference tracing records alphabetically on the sorted Subject Cross-Reference Tape.

Function 16: UPDATE MARC AND A/T FILES

1. The records from the New MARC Tape were merged with the records from the MARC Master Tape and these merged records became File 1 of the MARC master tape. In this process, both new records and revisions to previously entered records were included on the output tape. New records that were not revisions but which duplicated previously entered records were rejected, as were revision records for which no matching MARC records could be found. In both cases each rejected record was identified by an error message.

2. The records from the sorted Author/Title Tape were merged with the records from File 2 of the MARC Master Tape, and these merged records became File 2 of the MARC Master Tape. This process was identical to that used above for File 1; the same error identification and rejection techniques were used.

Function 17: UPDATE CROSS-REFERENCE FILES

1. The records from the sorted Subject Cross Reference Tape were merged with the records from File 3 of the MARC Master Tape. Only those records entered during the current week and the previous week were recorded in File 3. Records entered before this two-week period were deleted and did not appear on the output tape.

2. The records from the sorted Descriptive Cross Reference Tape were merged with the records from File 4 of the MARC Master Tape. Only those records entered during the current week and the previous week were recorded in File 4. Records entered before this two-week period were deleted and did not appear on the output tape. The resulting MARC Master Tape was a 9-track EBCDIC tape with both uppercase and lowercase character codes, containing all four files (MARC Records, Author/Title Records, Subject Cross-Reference Tracing Records, and Descriptive Cross-Reference Tracing Records) updated.

Function 18: LIST AND DUPLICATE MASTER TAPE

1. The MARC Master 9-track tape was converted to a 7-track BCD tape with upper and lower case character codes and both were duplicated to provide the necessary copies for the participants.

2. A sorted list of the LC catalog card numbers was printed for all of the records in File 1 of the MARC master tape. A copy of this list

was distributed to each participating library with a copy of the new Master Tape.

3. An alphabetical list of the Author/Title records added to File 2 of the MARC Master Tape during the current and previous week was printed and distributed to each participant.

The Weekly Processing Cycle responded to both the daily and weekly indicators described previously. The Daily Status Code (DSC) was used to identify the records on the MARC Work Tape that were ready to be added to the master files. The Weekly Status Code (WSC) was used to reflect the age of records on the MARC Master Tape.

Table 7 summarizes the action taken in the weekly processing functions based on the Daily and Weekly Status Codes.

TABLE 7.—*Weekly Action Taken in Response to Status Code*

		L (daily)	G (daily)	R	N	O	"blank"
Function 11a Separate Input Records		Place Rec. on MARC Work Tape Residual. Print L.C. & Control Numbers	Accept Rec. into Weekly Processor	No Change	No Change	N/A	N/A
Function 11b List Verified L.C. Numbers		N/A	List L.C. Catalog No. of MARC Rec.	No Change	No Change	N/A	N/A
Function 16 Update MARC & A/T Files	Rec. from MARC Work Tape	N/A	N/A	Replace Master Tape Rec. with New Record. Retain Status Code R.	Merge New Rec. onto Master Retain Status Code N.	N/A	N/A
	Rec. from MARC Mstr. Tape	N/A	N/A	Change Status Code from R to O.	Change Status Code from N to O.	Change Status Code from O to "blank"	Retain Status Code "blank"
Function 17 Update X-Ref Files	Rec. from MARC Work Tape	N/A	N/A	Replace Master Tape Rec. with New Record. Retain Status Code R.	Merge New Rec. onto Master Tape. Retain Status Code N.	N/A	N/A
	Rec. from Mstr. Tape	N/A	N/A	Change Status Code from R to O.	Change Status Code from N to O.	Delete Record from Master Tape.	Delete Record from Master Tape.

(The remaining functions do not utilize or change any status indicator codes.)

MARC INTERIM SYSTEM

The computer software described in the previous section was used to process MARC I records until October 1967, when the MARC Interim System was put into operation.

The MARC Pilot System was designed for the System/360 Model 30 with 16,000 core positions of memory. This was later augmented with additional module of 16,000 core positions and two disks. The original limitation of 16,000 positions of memory influenced the design of many functions of the MARC Pilot System, causing excessive amounts of administrative and technical overhead of the operation.

File maintenance capabilities were inexpedient and the system was not conducive to expansion. What was needed at the Library of Congress was an information system to process MARC records for distribution that, in addition, was capable of handling related bibliographic projects for the Library of Congress. The design of such a system had to be generalized and modular. Each module had to be capable of expansion for particular requirements without having the effect felt throughout the entire framework of the system. Both the MARC Interim System and MARC System are based on this modular concept. Since the programs in both systems are so interrelated, the

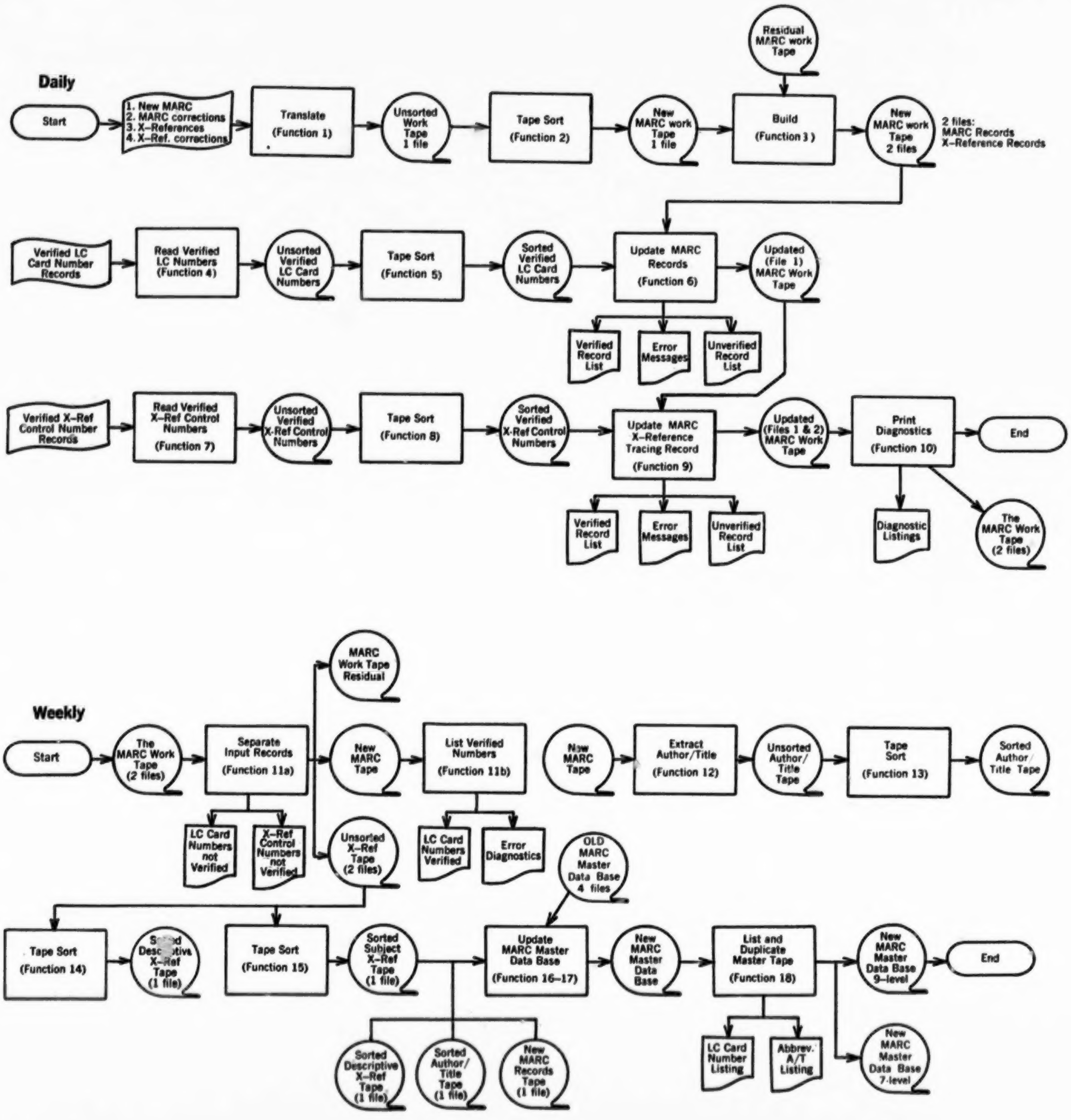

Daily
Start
1. New MARC
2. MARC corrections
3. X-References
4. X-Ref. corrections
Translate (Function 1)
Unsorted Work Tape 1 file
Tape Sort (Function 2)
New MARC work Tape 1 file
Residual MARC work Tape
Build (Function 3)
New MARC work Tape 2 files
2 files: MARC Records X-Reference Records
Verified LC Card Number Records
Read Verified LC Numbers (Function 4)
Unsorted Verified LC Card Numbers
Tape Sort (Function 5)
Sorted Verified LC Card Numbers
Update MARC Records (Function 6)
Updated (File 1) MARC Work Tape
Verified Record List
Error Messages
Unverified Record List
Verified X-Ref Control Number Records
Read Verified X-Ref Control Numbers (Function 7)
Unsorted Verified X-Ref Control Numbers
Tape Sort (Function 8)
Sorted Verified X-Ref Control Numbers
Update MARC X-Reference Tracing Record (Function 9)
Updated (Files 1 & 2) MARC Work Tape
Print Diagnostics (Function 10)
End
Verified Record List
Error Messages
Unverified Record List
Diagnostic Listings
The MARC Work Tape (2 files)
Weekly
Start
The MARC Work Tape (2 files)
Separate Input Records (Function 11a)
MARC Work Tape Residual
New MARC Tape
List Verified Numbers (Function 11b)
New MARC Tape
Extract Author/Title (Function 12)
Unsorted Author/Title Tape
Tape Sort (Function 13)
Sorted Author/Title Tape
LC Card Numbers not Verified
X-Ref Control Numbers not Verified
Unsorted X-Ref Tape (2 files)
LC Card Numbers Verified
Error Diagnostics
OLD MARC Master Data Base 4 files
Tape Sort (Function 14)
Sorted Descriptive X-Ref Tape (1 file)
Tape Sort (Function 15)
Sorted Subject X-Ref Tape (1 file)
Update MARC Master Data Base (Function 16–17)
New MARC Master Data Base
List and Duplicate Master Tape (Function 18)
New MARC Master Data Base 9-level
End
Sorted Descriptive X-Ref Tape (1 file)
Sorted Author/Title Tape (1 file)
New MARC Records Tape (1 file)
LC Card Number Listing
Abbrev. A/T Listing
New MARC Master Data Base 7-level

FIGURE 16.—*MARC Pilot System*

following section will describe them in terms of the MARC System. Thus the references are to data elements in the MARC II format.

The MARC System is divided into four major subsystems:

1. The Input Subsystem is concerned with the input, conversion, and formatting of bibliographic information into digital form from any medium, i.e., punched paper tape, punched cards, on-line terminals, etc.
2. The Storage Subsystem is concerned with the storage and maintenance of bibliographic records in digital form.
3. The Retrieval Subsystem is concerned with the language of retrieval, i.e., query, and the retrieval of records or parts of records or both from storage.
5. The Output Subsystem is concerned with the arrangement, the display as script in any form, e.g., printed, cathode ray tube, etc., and the transmission of records for distribution, e.g., magnetic tape, data transmission, etc.

Since the emphasis has been placed on the development of the MARC Distribution Service, the full potential of the MARC System has not yet been achieved. For example, the retrieval aspects are minimal. However, the framework has been provided upon which to build an increasingly sophisticated information system. The necessary interfaces between subsystems are common across the board, but the design of programs within each subsystem is independent of every other subsystem. There is evidence that, as the system expands and technology advances, the four subsystems will contract to two, namely, Input/Output and Storage and Retrieval.

The following subsections describe the system in terms of the functional operations performed during each subsystem and describe the flow of information through the individual component programs.

A. *Input Subsystem*

PRE-EDIT Program

This program is designed to accept punched paper tape input data encoded in any defined code set to process any length data field, any length identifying tag preceding the data field, any number of data fields, and data fields given in any order. The Dura Mach 10 code set is in current use. Each record is written on magnetic tape (Pre-edit Output Tape) in the sequence that it was read from paper tape, field by field. Therefore, the PRE-EDIT program can accommodate data from any application and is in no way restricted to the requirements of a defined format.

The processing of each field includes the following operations:

1. Performs any corrective actions indicated by a backspace code.
2. Deletes entire words when indicated by a delete code.
3. Deletes entire fields when indicated by a delete code.
4. Converts characters from one of two sets, each containing two cases, into single characters in EBCDIC code for System/360. Set One is the standard MARC keyboard (see chapter 7), and Set Two contains 88 additional characters. With the use of Set One and Set Two, the keyboard has the capacity for a total of 176 characters. The set shifts operate in the same manner as the uppercase and the lowercase shifts, i.e., one set shift negating the previous set shift. For example, punching the Set Two shift, then three characters, then the Set One shift, then two characters, results in three characters in Set Two mode followed by two characters in Set One mode.
5. Generates block terminators (special characters following the tag), field terminators, and record terminators.

FORMAT EDIT Program

This program reads the fields from magnetic tape as written by the PRE-EDIT program until a record separator is detected. The processing of each record results in the assembly of a logical record. The following operations are performed:

1. Variable fields are formatted, indicators are added, field terminators are generated, and the directory is constructed.
2. Each entry in the directory is augmented with an indicator describing the field's order in the input sequence.

The output to this program is a tape file of formatted records.

CONTENT EDIT Program

This program is the first in the series that is restricted, in that its output is a particular form of material in the MARC format. The processing of each record performs the following operations:

1. Fixed fields are formatted and validated. Numeric fields are checked for all numeric characters, alpha fields are checked for all alpha characters, etc.
2. Variable field tags are validated.
3. Record control number is validated.
4. Variable fields are scanned for keywords and certain fixed fields set by computer.

In addition to producing a magnetic tape of edited records in the MARC format, this program also produces the diagnostic used by the MARC editor (Figure 10). Each record is printed in a format similar to the format used when the record was keyed. As each data field is printed, all errors detected by the computer validation routines are printed. Depending on the severity of the particular error, a data field is left intact and so indicated, or a data field is deleted and so indicated, or in the most severe error conditions, an entire record is deleted and so indicated.

The programs described above operate on each type of record in the same manner, e.g., new MARC records, corrected MARC records, verified MARC records, or MARC records to be deleted.

SORT/MERGE Program

The MARC records are sorted by LC card number.

B. *Storage Subsystem*

UPDATE 1 Program

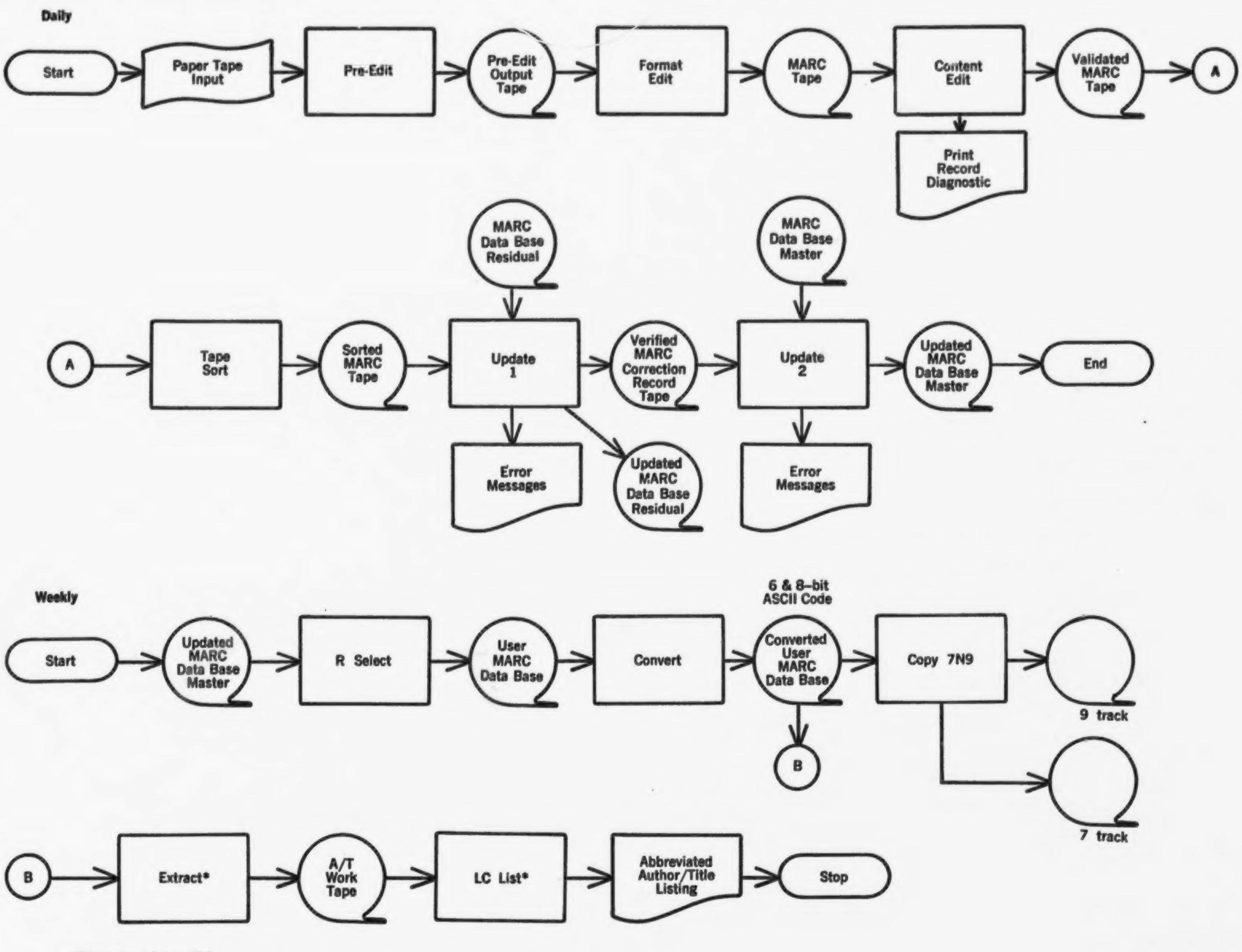
The UPDATE 1 Program performs initial file maintenance operations on the MARC Data Base Residual tape, i.e., the data base of records not yet declared error free. The following operations are performed:

1. Addition of entire records.
2. Deletion of entire records.
3. Changes to fields in records (add, delete, or replace a field).
4. Alteration in the status of a record, e.g., a record is declared verified.

UPDATE I produces a tape file containing verified records extracted from the MARC Data Base Residual. Another file is produced containing all remaining records. This second file serves as the updated MARC Data Base Residual for the next execution of the program. The program also produces a printed list of errors, impossible requests, etc.

UPDATE 2 Program

The UPDATE 2 Program performs file maintenance operations on the MARC Data Base Master tapes containing the verified records. The operations performed on this file are identical to those described in the UPDATE 1 Program, i.e., addition and deletion of records, addition, deletion, or replacement of fields. This


Daily
Start
Paper Tape Input
Pre-Edit
Pre-Edit Output Tape
Format Edit
MARC Tape
Content Edit
Validated MARC Tape
A
Print Record Diagnostic
MARC Data Base Residual
MARC Data Base Master
A
Tape Sort
Sorted MARC Tape
Update 1
Verified MARC Correction Record Tape
Update 2
Updated MARC Data Base Master
End
Error Messages
Updated MARC Data Base Residual
Error Messages
Weekly
Start
Updated MARC Data Base Master
R Select
User MARC Data Base
Convert
6 & 8-bit ASCII Code
Converted User MARC Data Base
Copy 7N9
9 track
B
7 track
B
Extract*
A/T Work Tape
LC List*
Abbreviated Author/Title Listing
Stop

*These procedures will be eliminated in the MARC Distribution Service.

FIGURE 17.—*MARC System Flow*

program also produces a printed list of errors, impossible requests, etc.

PRINT INDEX Program

The PRINT INDEX Program produces a listing of the LC catalog card numbers for the records being stored in the MARC Data Base.

C. *Retrieval Subsystem*

RSELECT Program

The RSELECT Program is a generalized select program, designed to perform the following operations:

1. Select specific records from the MARC Data Base.
2. Replace specified fixed fields with specified values.
3. Examine the contents of variable fields for specified values and select on the fields accordingly.
4. Compute the maximum, minimum, and average lengths of variable fields.

The output of RSELECT is a magnetic tape containing a full record format, or a partial record format as determined by use of control cards. In addition, the option exists to stipulate whether the selected records are to be written separately on one tape and records not meeting the criteria on another tape or in the case of 2 above, the specified records are to be updated, but all records are to be written as one file.

D. *Output Subsystem*

CONVERT Program

The CONVERT Program converts data from the internal LC processing format to the MARC communications format and translates the EBCDIC characters into ASCII 6-bit and 8-bit code.

COPY7N9 Program

The COPY7N9 Program duplicates 7-level tape and 9-level tape for distribution. These tapes are called User MARC Data Base Tapes.

EXTRACT Program (To be eliminated when MARC Distribution Service begins.)

The EXTRACT Program formatted an Author/Title record for each new and revised record on the User MARC Data Base Tape. These records are written on the Author/Title Work Tape.

LCLIST Program (To be eliminated when the MARC Distribution Service begins.)

The LCLIST Program prints an Abbreviated Author/Title Listing from the Author/Title Work Tape.

Programs in Progress

A number of new programs are also in progress which will expand the capabilities of the system. New programs in the Output Subsystem will include the following:

LIBRARY SORT Program

In June 1965, ISS Planning Memorandum No. 3 [3] indicated that a study of the filing rule problem would be made and a second report issued from the Library of Congress. Unfortunately, due to staffing problems, the filing rule study was not begun in 1965 but two years later in 1967.

A two-pronged attack was launched—a study of the filing rules themselves and a study of the methods of sorting MARC records by computer. In order to solve immediate needs within the Library of Congress and, in addition, provide a vehicle for testing sorting algorithms with the aid of the computer, a multiphase sort program was conceptualized that used the IBM System/360 DOS sort generator as the basic sort program. Each phase was designed to add sophistication to the preceding phase. Simultaneously an investigation was made of the need for writing a library sort program if the manufacturer's products would not suffice. At present, Phase I is completed and consists of two programs:

1. SKED. The selection of sort fields is controlled by parameter cards. Present capability permits the choice of fixed or variable fields or subfields within variable fields. SKED is a program that performs two basic functions before executing the IBM System/360 DOS Sort Generation.
   a. Builds a sort key to be used by the manufacturer's sort program.
   b. Duplicates each record the number of times necessary as determined by data in each record specified by parameter cards. For example, if the parameter card calls for a sort on the added entry variable fields, the record will be duplicated for each entry in the original record. The only difference in the duplicated record

will be the sort key, i.e., each sort key will reflect the information in the particular added entry.

Each field chosen for the sort key is translated to a sorting alphabet as opposed to a printing alphabet and the fields are physically placed next to each other, each field separated by a character smaller than any other permissible character.* The length of the sort key in Phase I is 256 characters.

2. POSTPASS. The POSTPASS Program is a modification of the IBM System/360 sort generator. The sort key is removed from the record before further processing.

BIBLIST Program

The BIBLIST Program is a generalized print program to format and print MARC records. The program is controlled by parameter cards which provide the following options:

1. Print selected records, i.e., by LC classification number, LC card number, author, title, subject headings, etc.
2. Format and print data fields within the record in any prescribed sequence.
3. Supply additional data than that already in the record.
4. Select special formats, e.g., hanging indention.
5. Select special format display, i.e., all capital letters, etc.
6. Select line length, margins, columns, page length, etc.

---

* This technique is used to prevent erroneous sorts based on the comparison of characters in unlike fields.

Support Programs for the Participating Libraries

To accelerate the use of the MARC data by the participating libraries, the Library of Congress provided a series of basic computer programs. In view of the time scheduled for the planning and implementation of the pilot project, this was a very ambitious undertaking, but it had the advantage of allowing the individual libraries to devote their efforts to developing special programs.

The Library recognized that it could not write programs for a wide variety of computers, so the decision was made to program for only two types of equipment.

To ensure that all participants could use the MARC tapes, the limitations imposed by the minimal computer configuration had to be considered. A survey was made by the contractor of each library's equipment and the two configurations selected, based on the results, were the IBM 1401 with 8,000 characters of core storage and none of the optional special features of this series, and the System/360 Model 30 with 16,000 characters of core storage. The selection of the limited 1401 configuration later proved to have been unnecessary and unfortunate since each 1401 installation did in fact have the special features. The report that one of the participating libraries had a 1401 with no special features was erroneous.

To the extent that it was possible to duplicate computer instructions for two distinctly different types of equipment, the functions performed by the programs for the System/360 and 1401 computers were identical, and the formats of the printed outputs were similar. All the programs had the capability of printing uppercase and lowercase characters with diacritical marks.

The function and the output format of each program is described below.

Bibliographic Listing Program

This program was designed to present a printout of the entire content of selected MARC records. Figure 18 illustrates the format in which these data were presented.

Each record was printed on a separate page, with the fixed field information appearing as words or abbreviations at the top of the page. For example, an A appearing in the field for type of main entry resulted in the printing of "Personal author" and an X in the bibliography field in the record appeared as a "Yes" on the listing. A blank in the field was printed "No."

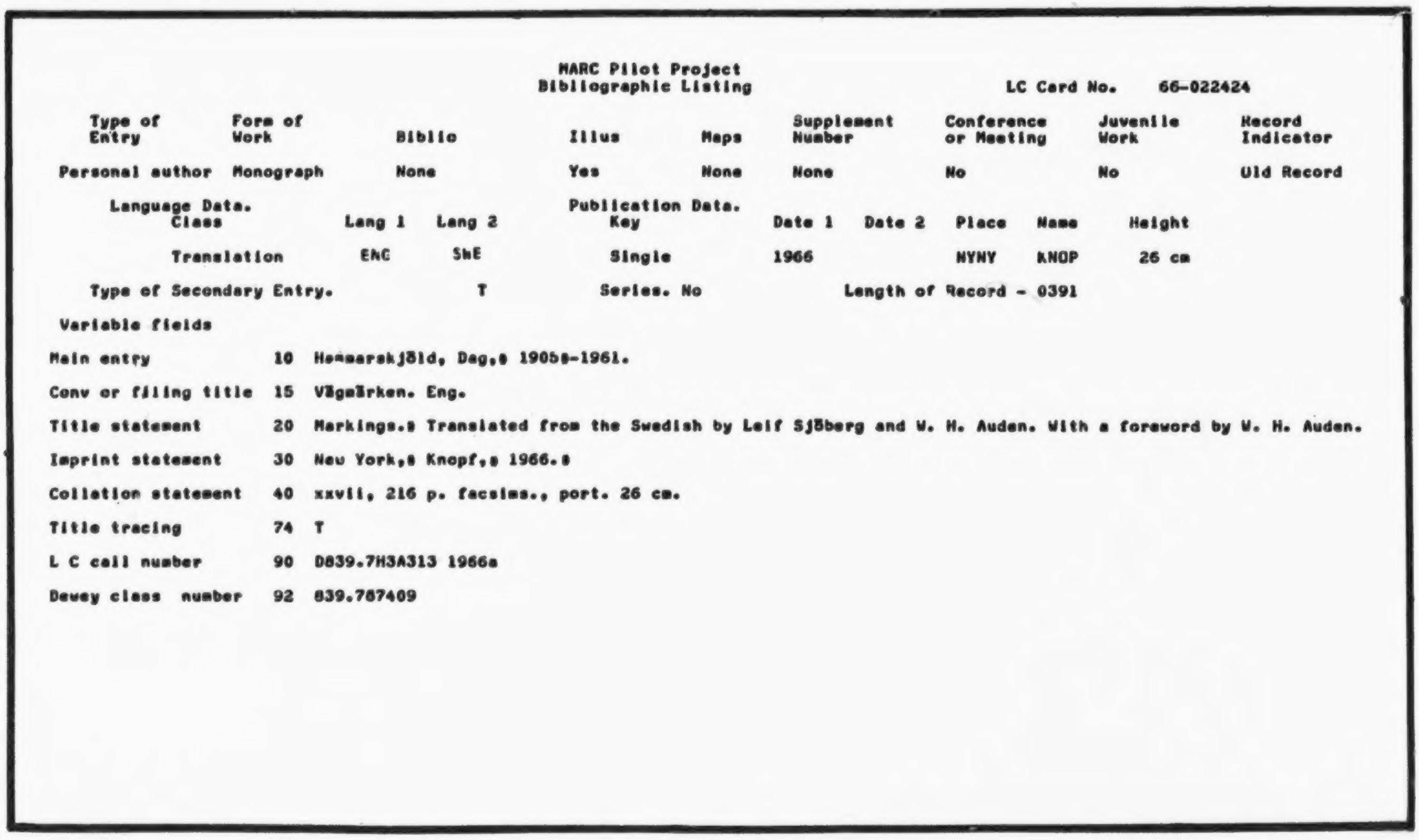
The variable fields of the MARC record were printed in the listing, with each field identified by its name and tag number. The total content of each field was printed in continuous form with no hyphenation at the end of the lines.

The first character of the content of the title tracing field (tag 74) contained the letter T. If only this letter appeared in this field, it indicated that the title tracing was to be taken from the title statement (tag 20) for overprinting 3x5 cards.

In the listing of the series tracing fields (tag 75), the first character was either an A or a T to signify a series author/title tracing or a series title tracing, respectively.

3 x 5 Catalog Card Program

This program was designed to present the bibliographic content of selected MARC records in the form of 3x5 cards that approximated the format of the Library of Congress typeset cards. The use of currently available computer systems to print these cards did not allow the flexibility that is available with the use of typeset cards. The equipment available printed 10 characters to the inch horizontally and 6 print lines to the inch vertically, which allowed a maximum of 40 characters per line and 17 lines per card. In addition, the character spacing was not variable as it is on the typeset card. Therefore a number of deviations from the tradi-



```
                                               MARC Pilot Project
                                              Bibliographic Listing                          LC Card No.    66-022424

     Type of       Form of                                                       Supplement     Conference     Juvenile     Record
     Entry         Work            Biblio           Illus        Maps            Number         or Meeting     Work         Indicator

  Personal author  Monograph       None             Yes          None            None           No             No           Uld Record

       Language Data.                               Publication Data.
            Class          Lang 1   Lang 2             Key               Date 1   Date 2   Place   Name     Height

            Translation     ENG      SWE               Single            1966              NYNY    KNOP      26 cm

     Type of Secondary Entry.         T               Series. No              Length of Record - 0391

  Variable fields

 Main entry             10  Hammarskjöld, Dag,‡ 1905‡-1961.

 Conv or filing title   15  Vägmärken. Eng.

 Title statement        20  Markings.‡ Translated from the Swedish by Leif Sjöberg and W. H. Auden. With a foreword by W. H. Auden.

 Imprint statement      30  New York,‡ Knopf,‡ 1966.‡

 Collation statement    40  xxvii, 216 p. facsims., port. 26 cm.

 Title tracing          74  T

 L C call number        90  D839.7H3A313 1966a

 Dewey class  number    92  839.787409
```

FIGURE 18.—*Participants' Bibliographic Listing*

tional catalog card were required. An example of the 3x5 computer-printed catalog card is shown in Figure 19.

GENERAL FORMAT

All catalog cards, including the main entry card and all added entry cards were printed in the general format described below. The specific format for each card type, i.e., main entry and extension cards, is described in the subsequent paragraphs.

On all cards, the main entry began in the 6th character position. All subsequent entries or continuations to the main entry began in the 8th character position, with an additional two-character indentation (to the 10th character position) for all new paragraphs. This allowed for a four-character left margin.

There was no right justification of card content and no hyphenation of words. All information was in complete words; thus there was a noticeable variation in the right margin from line to line.

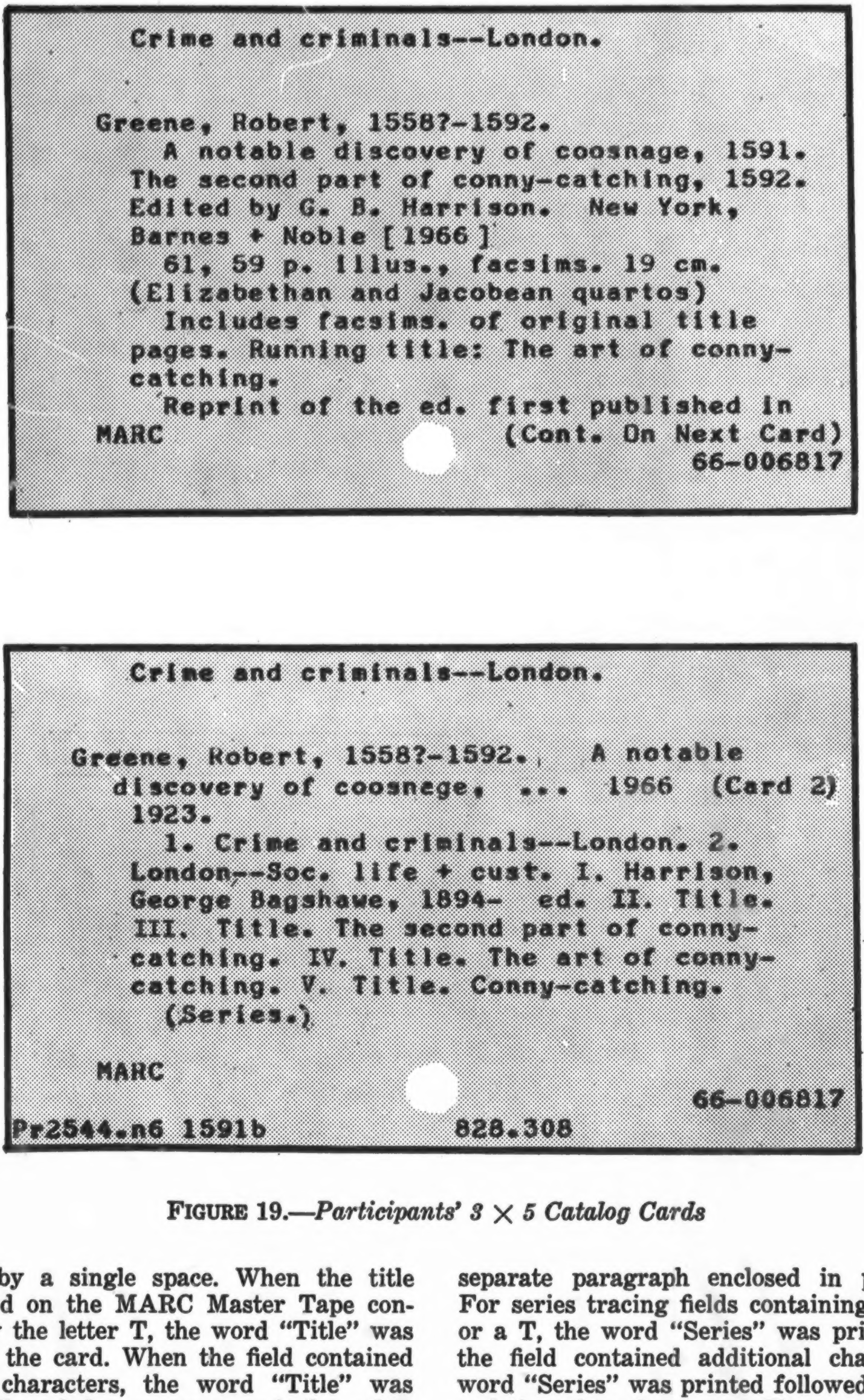
MAIN ENTRY CARD

The format of the main entry card provided the basis for all catalog cards produced. Tracing and extension cards exhibited only slight variations from this basic format. The following description relates to a single main entry card with no extensions.

Bibliographic information began with the main entry on line four and continued until complete, with no separation between lines unless a continuation card was required. The title paragraph followed the main entry, beginning in the 10th position of the next line and consisting of the title, edition, and imprint statements with a two-space separation between statements.

The collation, each series note, and each bibliographic note followed as separate paragraphs, with each paragraph beginning in the 10th position. (The first series note was enclosed in parentheses.)

Subject headings and secondary entries were grouped into another paragraph, identified by the Roman or Arabic numbering system and



```
    Crime and criminals--London.

Greene, Robert, 1558?-1592.
    A notable discovery of coosnage, 1591.
  The second part of conny-catching, 1592.
  Edited by G. B. Harrison.  New York,
  Barnes + Noble [1966]
    61, 59 p. illus., facsims. 19 cm.
  (Elizabethan and Jacobean quartos)
    Includes facsims. of original title
  pages. Running title: The art of conny-
  catching.
    Reprint of the ed. first published in
MARC                      (Cont. On Next Card)
                                    66-006817
```

```
    Crime and criminals--London.

Greene, Robert, 1558?-1592.   A notable
  discovery of coosnege, ...  1966  (Card 2)
  1923.
    1. Crime and criminals--London. 2.
  London--Soc. life + cust. I. Harrison,
  George Bagshawe, 1894-  ed. II. Title.
  III. Title. The second part of conny-
  catching. IV. Title. The art of conny-
  catching. V. Title. Conny-catching.
    (Series.)

MARC
                                    66-006817
Pr2544.n6 1591b         828.308
```

FIGURE 19.—*Participants' 3 × 5 Catalog Cards*

separated by a single space. When the title tracing field on the MARC Master Tape contained only the letter T, the word "Title" was printed on the card. When the field contained additional characters, the word "Title" was printed followed by a space and the other characters.

Series secondary entries were printed as a separate paragraph enclosed in parentheses. For series tracing fields containing only an A or a T, the word "Series" was printed. When the field contained additional characters the word "Series" was printed followed by a space and the other characters. The copy statement, if present, followed as a separate paragraph.

The three bottom lines were reserved for control information. The acronym MARC was printed on line 15 beginning in character position 6. The national bibliography number, if any, began in the 8th character position of line 16. The LC catalog card number was right-justified on line 16, with any prefix left-justified in the three positions preceding the numeric portion. Line 17 contained the LC call number beginning in the 2d character position and the Dewey Decimal number beginning in position 27.

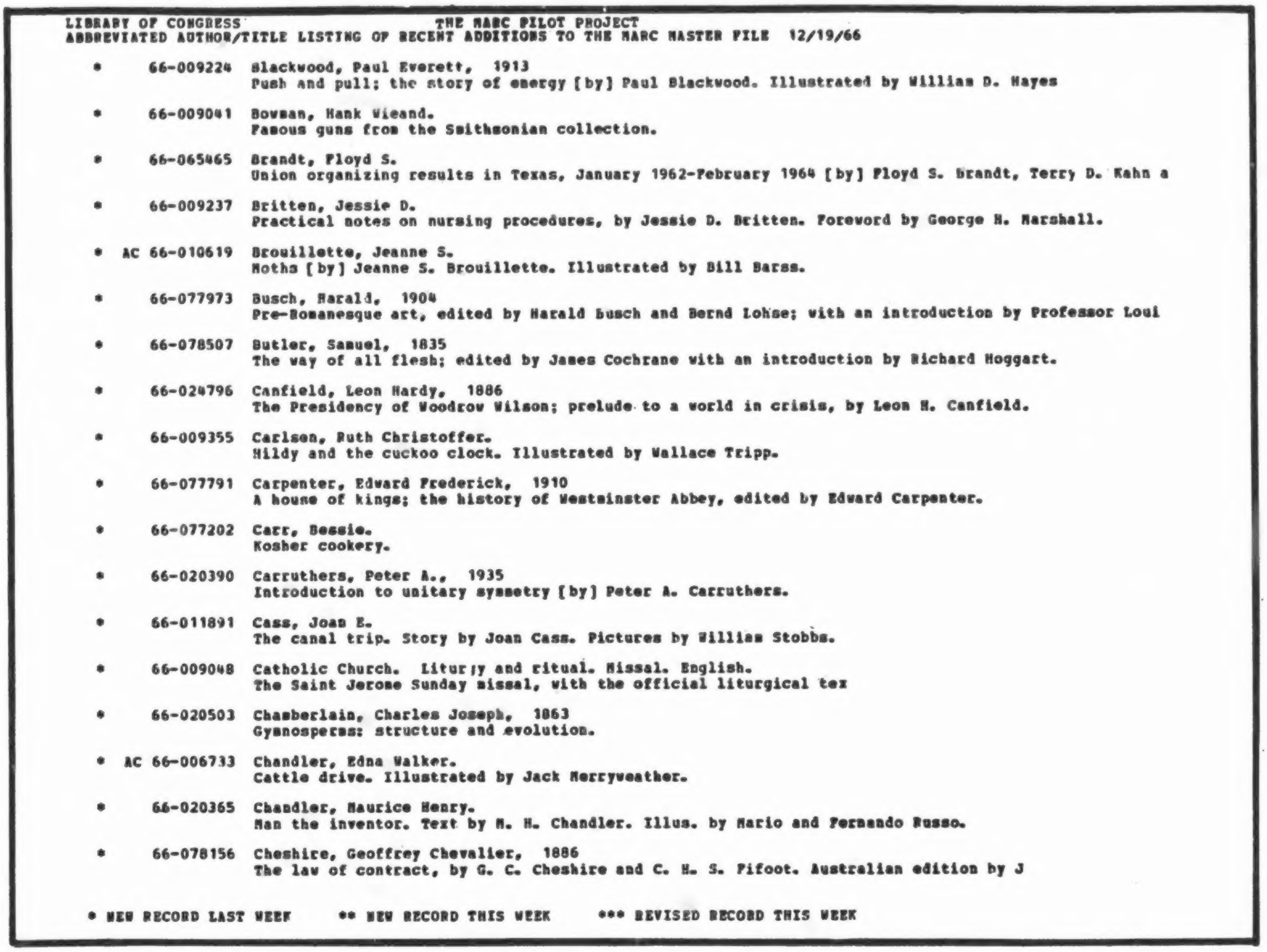
EXTENSION CARDS

When all the catalog information could not be printed on a single card, extension cards were used. The cards were arranged as follows:

1. Leader Card. The format of this card was identical to that of the main entry for the first 14 lines. Line 15 contained the acronym MARC, beginning in character position 6 and the phrase "(Cont. on next card)" right-justified. The LC catalog card number was right-justified on line 16.

2. Intermediate Extension Cards. A maximum of 66 characters of the main entry and title statement was printed, beginning in character position 6 of line 4 and continuing in character position 8 of line 5, if necessary. When the main entry exceeded one line, it was truncated and marked by ellipses at the end of the line. The title statement followed the main entry, separated by a single space. (When the main entry was truncated, the title statement appeared on line 5.) The date and card number also appeared on line 5, right-justified.

3. Last Extension Card. The first 14 lines of this card were identical in format to the inter-



```
LIBRARY OF CONGRESS                     THE MARC PILOT PROJECT
ABBREVIATED AUTHOR/TITLE LISTING OF RECENT ADDITIONS TO THE MARC MASTER FILE  12/19/66

   *      66-009224  Blackwood, Paul Everett,  1913
                     Push and pull; the story of energy [by] Paul Blackwood. Illustrated by William D. Hayes

   *      66-009041  Bowman, Hank Wieand.
                     Famous guns from the Smithsonian collection.

   *      66-065465  Brandt, Floyd S.
                     Union organizing results in Texas, January 1962-February 1964 [by] Floyd S. Brandt, Terry D. Kahn a

   *      66-009237  Britten, Jessie D.
                     Practical notes on nursing procedures, by Jessie D. Britten. Foreword by George H. Marshall.

   *   AC 66-010619  Brouillette, Jeanne S.
                     Moths [by] Jeanne S. Brouillette. Illustrated by Bill Barss.

   *      66-077973  Busch, Harald,  1904
                     Pre-Romanesque art, edited by Harald Busch and Bernd Lohse; with an introduction by Professor Loui

   *      66-078507  Butler, Samuel,  1835
                     The way of all flesh; edited by James Cochrane with an introduction by Richard Hoggart.

   *      66-024796  Canfield, Leon Hardy,  1886
                     The Presidency of Woodrow Wilson; prelude to a world in crisis, by Leon H. Canfield.

   *      66-009355  Carlsen, Ruth Christoffer.
                     Hildy and the cuckoo clock. Illustrated by Wallace Tripp.

   *      66-077791  Carpenter, Edward Frederick,  1910
                     A house of kings; the history of Westminster Abbey, edited by Edward Carpenter.

   *      66-077202  Carr, Bessie.
                     Kosher cookery.

   *      66-020390  Carruthers, Peter A.,  1935
                     Introduction to unitary symmetry [by] Peter A. Carruthers.

   *      66-011891  Cass, Joan E.
                     The canal trip. Story by Joan Cass. Pictures by William Stobbs.

   *      66-009048  Catholic Church.  Liturgy and ritual. Missal. English.
                     The Saint Jerome Sunday missal, with the official liturgical tex

   *      66-020503  Chamberlain, Charles Joseph,  1863
                     Gymnosperms: structure and evolution.

   *   AC 66-006733  Chandler, Edna Walker.
                     Cattle drive. Illustrated by Jack Merryweather.

   *      66-020365  Chandler, Maurice Henry.
                     Man the inventor. Text by M. H. Chandler. Illus. by Mario and Fernando Russo.

   *      66-078156  Cheshire, Geoffrey Chevalier,  1886
                     The law of contract, by G. C. Cheshire and C. H. S. Fifoot. Australian edition by J

 * NEW RECORD LAST WEEK     ** NEW RECORD THIS WEEK     *** REVISED RECORD THIS WEEK
```

FIGURE 20.—*Participants' Author/Title Listing*

mediate extension cards. Line 15 contained only the acronym MARC. Lines 16 and 17 contained the national bibliography number, LC catalog card number, LC call number, and Dewey Decimal number in the same positions they occupied on the main entry card.

ADDED ENTRY CARD

The added entry began in the 8th character position of line 1 and was continued, if necessary, on lines 2 and 3. The format for the remainder of the card was identical to that of the main entry and extension cards described above.

CONVENTIONAL OR FILING TITLE CARD

When a MARC record had a conventional or filing title, this title was printed on lines 1 and 2 beginning in position 12. (The larger left margin distinguished this type of card from an added entry card.) The remainder of the card format was the same as that of the main entry and extension cards.

MARC Call Card Program

This program was designed to produce MARC call cards for use with both the BIBLIOGRAPHIC LISTING and 3x5 CATALOG CARD programs. These call cards contained the LC catalog card numbers of records added to the MARC master tape during the current week and was used to select records for printing.

Abbreviated A/T Listing Program

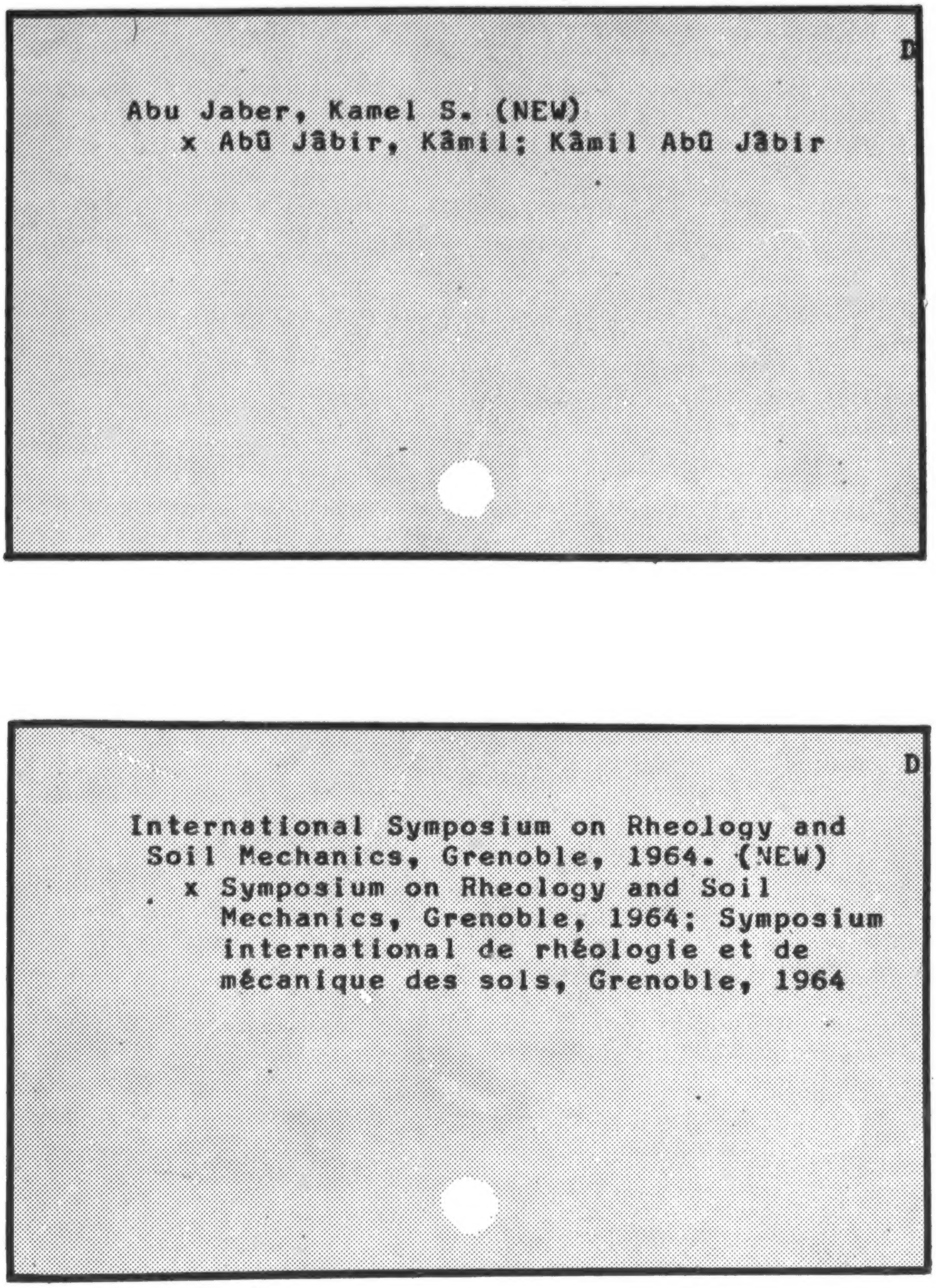
This program was designed to print out selected records or all of the records on File 2 of the MARC master tape in the form of a listing. The LC catalog card number and a maximum of 120 characters of information extracted from the author and title statements were included.

Figure 20 illustrates the format used in printing this author/title information.

3 x 5 Cross-Reference Card Program

This program was designed to print selected records or the entire content of Files 3 and 4 of the MARC master tape. The printing of these cards containing cross-reference tracing information followed the general format used for catalog cards. A maximum of 49 characters was printed on each line with a total of 17 lines per card. The character S or D was printed in character postion 49 on line 1 to indicate either a subject or descriptive cross-reference card, respectively. The cross-reference heading began in character position 6 on line 3. The action indicator followed the heading.

The cross-reference tracings followed the heading and action indicator. Each type of cross-reference tracing appeared in a separate paragraph starting in character position 9 for the *sa*, *xx*, and *x* (*see also*) references and in position 10 for the *x* references. When there was more than one reference of a given type, each reference was listed within the paragraph, separated by either a period or a semicolon. Examples of these cards are shown in Figure 21.



D

Abu Jaber, Kamel S. (NEW)
x Abû Jâbir, Kâmil; Kâmil Abû Jâbir

D

International Symposium on Rheology and Soil Mechanics, Grenoble, 1964. (NEW)
x Symposium on Rheology and Soil Mechanics, Grenoble, 1964; Symposium international de rhéologie et de mécanique des sols, Grenoble, 1964

FIGURE 21.—*Participants' 3 × 5 Cross-Reference Tracing Cards*

Codes Developed for Language, Publisher, and Place of Publication

During the time period from the publication of the ISS Planning Memorandum Number 3 to the design of the MARC Pilot System, there was much discussion about the usefulness of coding language, publisher, and place of publication. Some librarians felt all or some subsets of the three were critical to the identification of bibliographic material and therefore should be coded for rapid access in a machine system. Others argued that the information was already recorded in the variable field data and, although not as readily available, could be identified at input and made accessible for machine retrieval.

Since MARC was to be a pilot, the designers of the system considered it worthwhile to devise a set of codes and provide the means to test their utility in an operational environment. Any future MARC system could include or exclude these data elements based on empirical data gathered from an operational test.

There were many efforts in progress to design efficient coding schemes. The MARC staff investigated and evaluated as many of these as was possible during the short time allowed to accomplish the task. With the aid of the contractor, codes for language, publisher, and place of publication were developed to serve as the basis for the beginning operation. As the project progressed, new names were added to the list as needed.

The project was not underway for long before it became apparent that there were many difficulties involved in maintaining the publisher code and, consequently, it was a costly addition to the record. It was estimated that the cost of including this code exceeded five cents per record. This high cost was caused by research necessary to establish new codes.

It was necessary to gather sufficient information to distinguish a publisher from a firm with a similar name. Many trade publishers used more than one imprint, and in those cases a discrete code for each imprint was established. At times it was very difficult to determine if two codes for one imprint had been made or one code for two imprints. For example, considerable time was consumed determining that Macmillan of Canada; Macmillan, New York; Macmillan, London; and Macmillan, Melbourne, were two companies with branch offices: The Macmillan Co., New York, and Macmillan and Co., Ltd., London.

The publisher directories were only partially helpful in solving such problems, because none of them were complete, and different directories listed different forms of names for companies and imprints.

When the publisher code was devised, it was anticipated that, in time, the number of imprints requiring the establishing of new codes would diminish. However, after a year, no noticeable decrease could be seen. For that reason no new publisher codes were established after June 1967.

In addition to the high cost of maintaining the publisher code, it became apparent that little use was being made of it. When plans for the development of a standard book number became known, the potential value of the code diminished. Because of the high cost of the code and its doubtful value, it will be dropped with the implementation of the MARC II format.

The code for place of publication was also difficult to maintain. It was anticipated that the number of places (towns or cities) requiring codes would diminish in time, but they did not. The amount of research required to determine the correct spelling and location of small towns, both in the United States and abroad, was both time consuming and costly.

The consensus from the evaluation of the MARC I format indicated that a place code was useful but that a code for country of publication was sufficient and city of publication should be discontinued. Therefore, the place codes for the MARC II format will not include cities.

The base list of languages, with relatively few additions, has served the MARC system adequately. The evaluation of the MARC I format, however, indicated that the fourth character of the language code (representing language group) was of doubtful value and should be dropped. Languages in the MARC II format will be represented by three-character mnemonic codes.

The codes devised for the MARC Pilot System are being reevaluated in conjunction with the other two national libraries. The staff members responsible for the project are surveying work in progress in this area in both the Federal and the private sector in the hope of arriving at a system of encoding language and country that will be acceptable as a standard.

The following section describes the codes used in the pilot project. Up to the time of this report, codes had been assigned to 1,753 publishers, 1,402 cities, 288 countries, and 493 languages.

Publisher Code

The code used for the MARC Pilot Project to represent a publisher in the fixed field was a two-, three-, or four-letter mnemonic code. To keep the task of establishing codes for publishers within practical limits, it was necessary to restrict the assigning of discrete codes to certain categories of publishers. The following guidelines were used:

1. Discrete codes were assigned to:
    a. Trade publishers
    b. Universities and colleges
    c. Government printers
2. Four different kinds of general codes were used for government agencies in the United States at various levels, i.e., Federal, State, county, and city.
3. General codes were used for agencies of foreign governments without distinction as to level of government.
4. One code was used to indicate that a publisher was known, either from the imprint or from the main entry, but that the publisher was not being assigned a discrete code.
5. One code was used to indicate that the publisher was not known. The assignment of publisher codes was based on the publisher's statement appearing in the imprint. If the publisher's statement was omitted because the same corporate body was responsible both for authorship and publication, the publisher's code was assigned from the main entry.

Trade publishers were defined as businesses or organizations whose primary purpose was publishing. University presses were included in this category. In case of doubt as to whether a particular publisher was a trade publisher, it was necessary to consult lists of publishers in standard reference sources. In addition, the R. R. Bowker Company allowed the Library to use their coded master list of publishers. When a new trade publisher did not appear on any list of publishers consulted, a discrete code was not used until that publisher had appeared as a publisher on at least five records in the MARC pilot project. The original list was based on publishers included in the following directories:

1. Literary market place. 1963/64 ed. New York, R. R. Bowker Co. [1963]
2. American book trade directory. 17th ed., 1965–66. New York, R. R. Bowker Co., 1965.
3. The British national bibliography. 1963 annual vol. London, Council of the British National Bibliography, 1964.
4. Publishers' international year book. 2d ed. London, A. P. Wales, 1962.

The form of the publisher's name was verified for accuracy and completeness in standard reference sources. The name of the publisher was followed by the city in which the publisher was located. The name of a publisher as it appeared on a catalog card was not necessarily in the same form as the full name given in the MARC project list. The fuller name was given in the list on the assumption that it was easier to distinguish between publishers with similar names. The term "publisher" was used to include publisher's imprints (Clarendon Press, imprint of Oxford University Press) and divisions of publishing houses (Gregg Division, McGraw-Hill Book Company).

Names of publishers that were names of persons were entered in the list under the surname (DAY, JOHN, CO., N.Y.), with cross references from the forename form. In order to prevent incorrect computer alphabetizing, all words in the major portions of publishers' names were written out. Hyphens in compound names were not included for the same reason. Terms appearing at the end of publishers' names such as Co., Inc., Ltd., etc. were abbreviated.

The code for trade publishers was the code used on the Bowker list if the publisher was represented there. If not, the publisher was assigned a code.

Discrete codes were assigned to universities and colleges. Works published by subdivisions of universities and colleges were assigned codes for the highest level only, i.e., the university or college itself, not the specific subdivision. This code was assigned on the basis of the imprint or the main entry if no publisher appeared in the imprint. University presses received their own discrete codes, distinct from the codes for the universities with which they were associated.

Discrete codes were assigned to government printers, e.g., Government Printing Office (GPO), Her Majesty's Stationery Office (HMSO), etc. When individuals were named as government printers, e.g., "R. Duhamel, Queen's printer and controller of stationery" or "J. R. Lee, Acting Govt. printer," an arbitrary code was used representing the category "government printer" rather than any named individual. Thus, for Canada, "Queen's printer" was coded CDGP, and any individual named in the imprint was ignored.

One code (USFZ) was used for all agencies of the Federal Government of the United States. The code was used when the name of a Federal agency appeared in the imprint as publisher or when the name of an agency appeared only in the main entry, indicating that the agency was both the author and publisher. If an agency's name appeared in an imprint as well as the Government Printing Office, the code USFZ was used; if only the Government Printing Office appeared in the imprint, the code GPO was used.

When agencies of State, county, or city governments were publishers, the two-letter code for the State, and the letter S for State, K for county, and C for city and the letter Z were used to make up the code. The letter Z was used as the fourth character to insure uniqueness of the codes. Thus, for government agencies in the State of Connecticut, CNSZ indicated a State agency, CNKZ a county, and CNCZ a city agency. If the entry took a form indicating a government but without a subdivision, and if there was no publisher in the imprint, the code indicating the appropriate level of government was used. For example, if the entry was "Connecticut" and there was no publisher in the imprint, the code used was CNSZ.

If a publisher was an agency of any foreign government except Great Britain, one code was used regardless of the level of government involved. The code used consisted of the two-letter code for the country and the letters SZ. However, if the catalog entry for the agency involved was "Gt. Brit.," the code used was GBSZ. This was necessary because England, Scotland, Ireland, and Wales were given separate codes. These codes, however, were still used with the letters SZ to represent a government agency at any level within those jurisdictions. Thus, if the Information Department of the city of Birmingham, England, published a work, the code used was ENSZ. If the Ministry of Labour of Great Britain published a work, the code was GBSZ.

The coding of a body as a government agency was confined to those bodies that were direct agencies of government and excluded "certain classes of institutions and other bodies created, maintained, controlled, or owned by governments" (ALA rule 72A(1)).

The code SBDY was used for publishers not in the categories discussed above. This code indicated that the publisher is known, either from the imprint or from the main entry. When the publisher was not known, the code ANON was used.

Place of Publication

The code devised for the MARC Pilot Project to represent place of publication in the fixed field was a four-letter mnemonic code. The first two letters of the code represented the country, if foreign, or the State within the United States (EN=England; AZ=Arizona). The last two letters represented the city. This system allowed a maximum of 676 cities within each country or State (USA) to be represented. The code for city was not unique unless combined with the code for country or State. Example: ENCA=Cambridge, England; MSCA=Cambridge, Massachusetts. The code XXXX was used if the place of publication was not known. If the country or State of publication was known but not the city, the country code was used in combination with XX.

The sources of information about place names and places of publication are listed below:

1. U.S. *General Services Administration. Office of Finance and Administration.* Geographical location codes. Washington [1965]
2. U.S. *Social Security Administration.* Geographic codes for State and country of residence. [Washington] 1964.

2

3. Literary market place. 1963/64 ed. New York, R. R. Bowker Co. [1963]

4. The British national bibliography. 1963 annual vol. London, Council of the British National Bibliography, 1964.

5. Publishers' international year book. 2d ed. London, A. P. Wales, 1962.

6. Bartholomew, John George. Gazetteer of the British Isles. 9th ed., reprinted, incorporating a summary of the 1961 census. Edinburgh, J. Bartholomew, 1963.

7. Rand, McNally and Company. Commercial atlas and marketing guide. 97th ed. Chicago [1966]

8. Rand, McNally and Company. Cosmopolitan world atlas. Chicago [1962]

9. National Geographic Society, *Washington, D. C. Cartographic Division.* National Geographic atlas of the world. Washington, 1963.

10. The Columbia Lippincott gazetteer of the world. Morningside Heights, New York, Columbia University Press, by arrangement with J. B. Lippincott Co., 1952.

11. Reader's Digest Association, Ltd. The Reader's Digest complete atlas of the British Isles. [1st ed.] London [1965]

12. U. S. *Department of State. Office of the Geographer.* Status of the world's nations. [Washington, Office of Media Services, Bureau of Public Affairs] 1962. (Its Geographic bulletin no. 2)

Language Code

To represent language in the fixed field a three- or four-letter mnemonic code was devised for the MARC Pilot Project. The first three letters of the code represented a discrete written language; the last letter, when used, represented a language grouping either by language family or geographical area. The first letter of the code was in all cases the first letter of the name of the language. The English name of a language was used when one existed. Where possible, the first three letters of the code were the first three letters of the name of the language.

An arrangement was devised to provide for the grouping of languages either by family or geographical area. Twenty-three groupings were established, each represented by a letter which, when used, was the fourth letter of the code for a particular language. In addition, the letter Y was used for languages not in one of the 23 groups. The codes assigned for major languages (primarily European) did not include a fourth letter for language grouping. This was necessary to prevent too many titles being recalled when performing a machine search by the language grouping code.

Only two languages could be represented in the fixed fields. However, it was possible to indicate more than two languages for a particular work by using a Z as the fourth character of the code for language 2. Thus, if a work was in English, French, and German, the codes ENG and FREZ were used. The Z in the code for French indicated that more than two languages were involved in describing the complete language characteristics of a work. Note that the use of the Z referred to number of languages and did not necessarily mean that the *same* text was in more than two languages. A work in English with summaries in French and German also took the codes ENG and FREZ. When a code used for language 2 consisted of four letters, a Z could be used to replace the fourth letter when applicable. Two codes used in special circumstances were UNK (Unknown) and MULZ (Multilingual).

The initial list of languages used for compiling the codes was obtained by asking various divisions of the Reference Department for lists of languages in the division's field of interest. Most of the languages were represented in the Library of Congress collections. Additional languages were added by MARC editors as the need arose. Spelling and form of name of each language was checked against a list compiled by the Center for Applied Linguistics, Washington, D.C., for the National Register of Scientific and Technical Personnel of the National Science Foundation. Staff members of the Center made available revisions and corrections to the Center's list. Those languages which could not be correlated with the Center's list were omitted.

The following procedures were used in constructing the list for languages:

1. The Bantu languages of Africa were listed and coded under the base name, not the prefix. For example, EciJita was found on the list under the base name Jita. Following the base name, the prefix and the base name were given, all enclosed within brackets:

[JITA ECIJITA]

Common prefixes were: Chi-, Ci, E-, Echi-, Eci-, Eki-, Esi-, Gi-, I-, Ichi-, Iki-, Ishi-, Isi-, Ki-, Ku-, Li-, Lu-, M-, Olu-, Oru-, Osi-, Se-, Si-, U-, Ulu-, Uru-, Usi-.

2. Some language names had variant spellings using several combinations of letters. This information was given after the form of name chosen for the list and was enclosed within brackets:

ACOLI [AC (H) O (O) LI]

The above name could be spelled using any combination of the letters within parentheses, i.e., ACOOLI, ACHOOLI, and ACHOLI.

3. Names following a particular entry on the list and separated from the entry by an "=" sign were variant names.

AGAU=AGAW

4. When more than one language had the same name, distinguishing information was added and enclosed within parentheses:

KOTA (INDONESIAN)
KOTA (AFRICAN)
KOTA (INDIA)

5. The signs /, //, !, and # were used to indicate clicks in the Bushman languages. Although these signs were always used in spelling the names of the languages, they were ignored in alphabetization. In order to give this information without disrupting machine sorting of the names, the languages in the list were given in a form without any signs followed by the form including signs and enclosed within brackets:

NGKE [//NG!KE]

Character Sets

To transcribe and print out bibliographic data, it was necessary to design a character set which would include the special characters and diacritical marks used in the major roman alphabet languages. Early in the planning phases of the project, attention was drawn to the work in progress by the Library Typewriter Keyboard Committee of the Resources and Technical Services Division of the American Library Association. This committee, chaired by C. Donald Cook, had surveyed a representative number of large research libraries to determine the optimum character set for a Selectric typewriter designed principally for cataloging but to include some characters for accounting, etc. The Library Typewriter keyboard is shown in Figure 22. In the interest of standardization and to facilitate the choice of a character set for MARC I, the Library decided to base the character set on the work done by Mr. Cook's committee. The MARC keyboard, designed in consultation with Mr. Cook and the IBM engineering staff, exhibited minor differences from the standard library keyboard primarily because it was oriented toward the preparation of data for input to electronic data processing equipment. In obtaining the graphics required for machine input, the MARC keyboard lost the haček and the upper case duplication of the comma, period, and hyphen, but added separate keys for zero (Ø) and one (1) as well as two graphics—# and $—to be used as delimiters. This keyboard was capable of representing most of the major western European languages (Table 8). The keyboard is illustrated in Figure 23.

A print train for bibliographic data was designed at the same time. Since speed was a factor, it was decided to limit the characters to those graphics that would not require any original artwork, i.e., those characters which had already been designed by the manufacturer. The resultant print train included all the characters on the MARC keyboard with two additional diacritical marks, the haček (ˇ) and the superior dot (˙). The 240-position print train was divided into two identical sets of 120 characters. Table 9 lists the characters on the train.

Since the participant libraries had equipment using 7-level tapes and 9-level tapes, it was necessary to represent the character set in both 6-bit and 8-bit codes. The MARC record for 9-level tape users was written in Extended Binary Coded Decimal Interchange Code (EBCDIC). Special characters and diacritics that were not part of the standard EBCDIC code set were assigned unused codes. The MARC record for 7-level tape was written in Binard Coded Decimal (BCD). Because of the limited number of codes in BCD, some of the characters were represented by two codes. All upper case characters and diacritics were represented by a character code preceded by the word-separator code. It was possible to condense the two codes into one internally by reading the magnetic tape in the "load" mode. Both 6-bit and 8-bit codes are shown in Table 10.

Although the data for the MARC pilot represented English language monographs, the catalog records frequently contained information in other languages and in some cases in nonroman

TABLE 8.—*Language Capacity of MARC I Keyboard*

ROMAN ALPHABET	NONROMAN (romanized form)
Afrikaans Albanian Dutch English Esperanto Estonian Finnish French German Hawaiian Italian Latin Portuguese Spanish Swedish Welsh	Greek
(with improvisation) Danish Norwegian Rumanian	(with improvisation) Chinese Japanese Korean

5-8/5-9

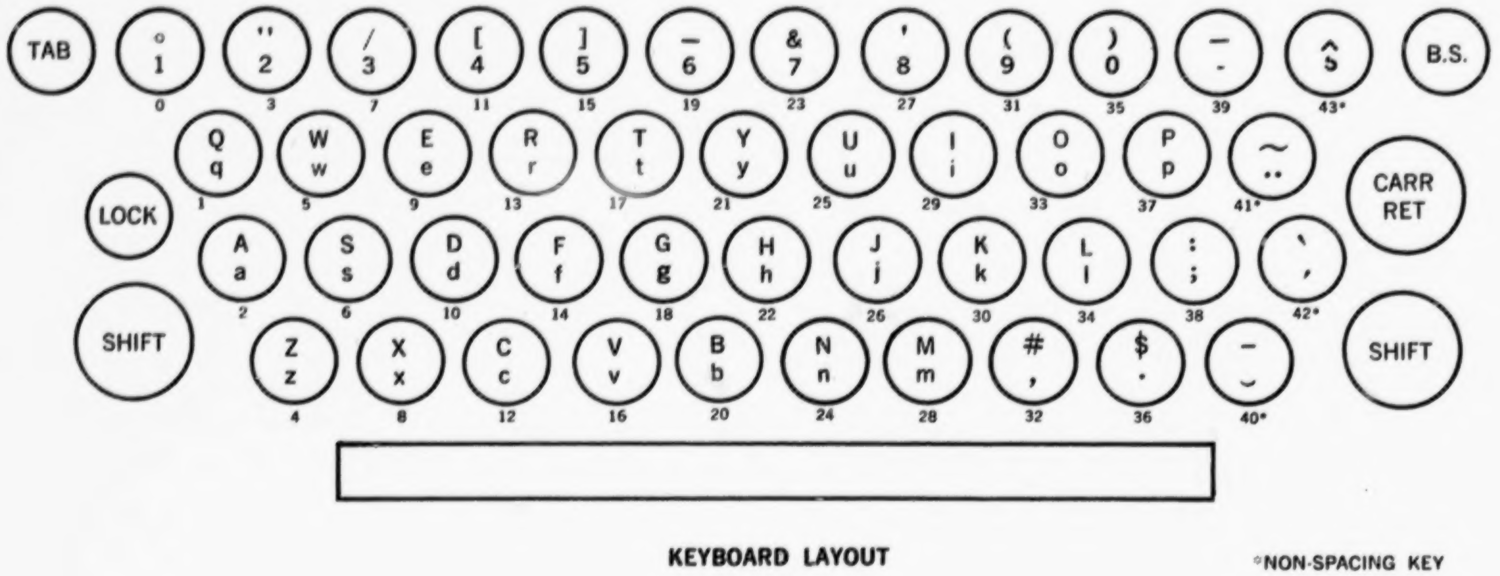


FIGURE 22.—*Library Typewriter Keyboard*

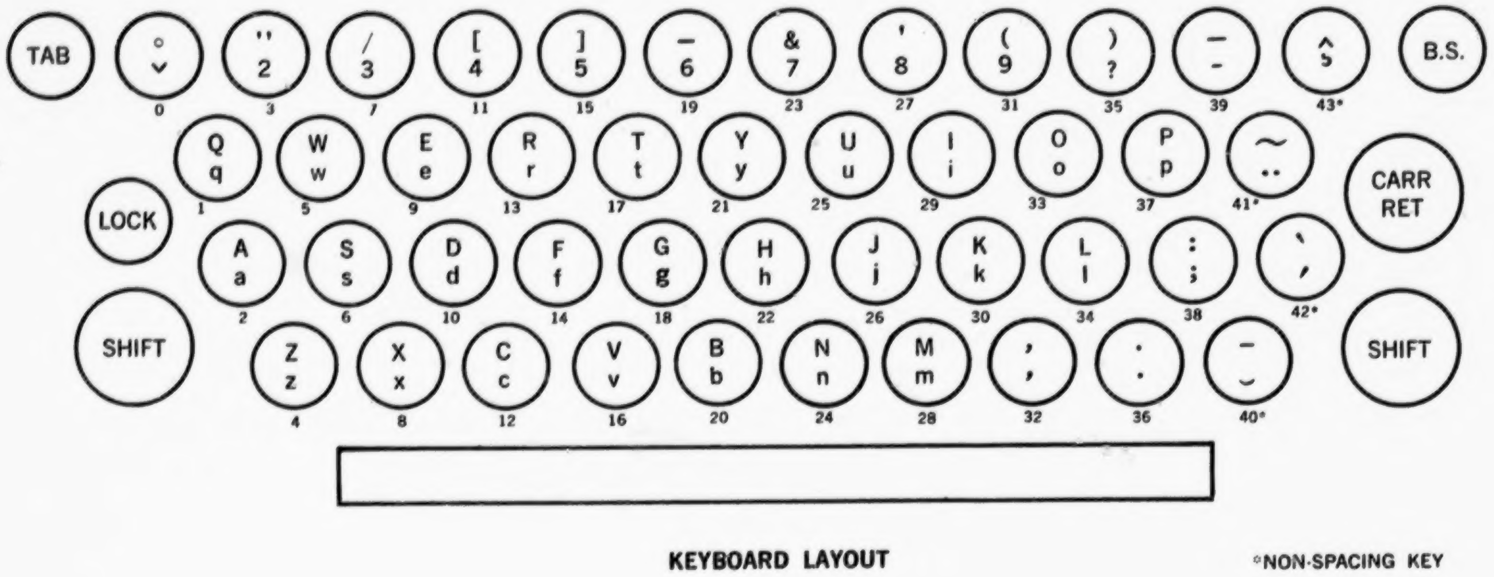


FIGURE 23.—*MARC Pilot Project Keyboard*

alphabets. Any record which could not be transcribed using the characters on the MARC keyboard was rejected. These records provided the basis for further study of the characters needed to represent bibliographic information.

From the beginning of the pilot, the Library recognized the importance of standardization of character sets. Once the pilot was actually put in operation, attention was turned to the development of an expanded set to serve all the major roman alphabet languages and the romanized forms of nonroman alphabets as well. The planning of the MARC Distribution Service included an in-depth analysis based on the pilot experience and on a study of the many languages represented on Library of Congress catalog cards. This investigation resulted in an extensive list of characters and diacriticals.[5]

This list was reviewed by the National Agricultural Library and the National Library of Medicine, and these libraries suggested other special characters needed to transcribe scientific and technical information. From this a composite character set was drawn up representing the needs of all three national libraries. The basic criteria used in the design of this set were frequency of occurrence of a character, necessity of expressing the character, and the possibility of substituting one for another or of expressing a character by writing it out.

Once the character set was designed, the next problem was how to represent these characters

TABLE 9.—*MARC Pilot Project Print Train*

PRINT POSITION	GRAPHIC	NAME
1	1	
2	2	
3	3	
4	4	
5	5	
6	6	
7	7	
8	8	
9	9	
10	0	
11	=	Equals
12	.	Period
13	/	Virgule
14	S	
15	T	
16	U	
17	V	
18	W	
19	X	
20	Y	
21	Z	
22	,	Comma
23	#	Pound sign
24	&	Ampersand
25	J	
26	K	
27	L	
28	M	
29	N	
30	O	
31	P	
32	Q	
33	R	
34	-	Hyphen
35	"	Double quote
36	:	Colon
37	A	
38	B	
39	C	
40	D	
41	E	
42	F	
43	G	
44	H	
45	I	
46	+	Plus
47	a	
48	b	
49	c	
50	d	
51	e	
52	f	
53	g	
54	h	
55	i	
56	j	
57	k	
58	l	
59	m	
60	n	
61	o	
62	p	
63	q	
64	r	
65	s	
66	t	
67	u	
68	v	
69	w	
70	x	
71	y	
72	z	
73	@	At sign
74	'	Apostrophe
75	?	Interrogation mark
76	;	Semicolon
77	±	Plus or minus
78	!	Exclamation
79	$	Dollar sign
80	*	Asterisk
81	%	Percent
82	^	Circumflex
83	`	Grave
84	´	Acute
85	~	Tilde
86	¸	Cedilla
87	¯	Macron
88	˘	Breve
89	˙	Superior dot
90	¨	Umlaut
91	ˇ	Hacek
92	ℓ	Script L
93	/	Polish L slash
94	c	
95	d	
96	e	
97	°	Angstrom
98	¬	Not
99	(	Left paren
100	)	Right paren
101	≠	Not equal
102	+	Plus
103	>	Greater than
104	<	Less than
105	≤	Less than or equal to
106	≥	Greater than or equal to
107	¢	Cent
108	□	Lozenge
109	[	Left bracket
110	]	Right bracket
111	}	Right brace
112	{	Left brace
113	⌊	Left sub-bracket
114	⌋	Right sub-bracket
115	r	
116	s	
117	t	
118	_	Underscore
119	-	Hyphen
120	|	Vertical line

in coded form. The technical specifications governing the coding were (1) the set had to be structured to facilitate derivation to a larger or a smaller code pattern; (2) the members of the character set had to be capable of being ordered in a pattern prescribed by popular usage, i.e., sorting order A-Z, 0-9, etc., and (3) the character set chosen had to be related to the present technology of Input/Output devices.

One possibility was to continue to use EBCDIC for 8-bit codes and BCD for 6-bit

TABLE 10.—*MARC Pilot Project Character Set*

NATURAL LANGUAGE SYMBOL	EBCDIC 8-BIT CODE (HEXA-DECIMAL)	BCD 6-BIT CODE	NATURAL LANGUAGE SYMBOL	EBCDIC 8-BIT CODE (HEXA-DECIMAL)	BCD 6-BIT CODE
space	40		±	9E	Not assigned
`	41	*WS*, 1	□	9F	Not assigned
´	42	*WS*, 2	/	A0	Not assigned
¯	43	*WS*, 21	s	A2	A 2
¨	44	*WS*, 4	t	A3	A 21
°	45	*WS*, 4 1	u	A4	A 4
~	46	*WS*, 84 1	v	A5	A 4 1
^	47	*WS*, 421	w	A6	A 42
˘	48	*WS*,BA	x	A7	A 421
¸	49	*WS*, 8 1	y	A8	A8
¢	4A	A	z	A9	A8 1
•	4B	BA8 21	⌊	AB	Not assigned
<	4C	BA842	⌊	AD	BA842
(	4D	BA84	≥	AE	Not assigned
+	4E	BA	ℓ	B0	Not assigned
|	4F	B 8421	•	B7	Not assigned
&	50	BA	∨	B9	Not assigned
!	5A	B 8 2	┘	BB	Not assigned
$	5B	B 8 21	┌	BD	B 8421
*	5C	B 84	≠	BE	Not assigned
)	5D	BA84 1		C0	BA8 2
;	5E	B 842	A	C1	*WS*,BA 1
	5F	B 8421	B	C2	*WS*,BA 2
-	60	B	C	C3	*WS*,BA 21
/	61	A 1	D	C4	*WS*,BA 4
,	6B	A8 21	E	C5	*WS*,BA 4
%	6C	A84	F	C6	*WS*,BA 42
_	6D	A842	G	C7	*WS*,BA 421
>	6E	842	H	C8	*WS*,BA8
?	6F	BA8 2	I	C9	*WS*,BA8 1
:	7A	B 84 1	J	D1	*WS*,B 1
#	7B	8 21	K	D2	*WS*,B 2
@	7C	Not assigned	L	D3	*WS*,B 21
'	7D	84	M	D4	*WS*,B 4
=	7E	842	N	D5	*WS*,B 4 1
"	7F	A8421	O	D6	*WS*,B 42
a	81	BA 1	P	D7	*WS*,B 421
b	82	BA 2	Q	D8	*WS*,B 8
c	83	BA 21	R	D9	*WS*,B 8 1
d	84	BA 4	S	E2	*WS*, A 2
e	85	BA 4 1	T	E3	*WS*, A 21
f	86	BA 42	U	E4	*WS*, A 4
g	87	BA 421	V	E5	*WS*, A 4 1
h	88	BA8	W	E6	*WS*, A 2
i	89	BA8 1	X	E7	*WS*, A 421
{	8B	Not assigned	Y	E8	*WS*, A8
≤	8C	Not assigned	Z	E9	*WS*, A8 1
j	91	B 1	0	F0	8 2
k	92	B 2	1	F1	1
l	93	B 21	2	F2	2
m	94	B 4	3	F3	21
n	95	B 4 1	4	F4	4
o	96	B 42	5	F5	4 1
p	97	B 421	6	F6	42
q	98	B8	7	F7	421
r	99	B8 1	8	F8	
}	9B	Not assigned	9	F9	8 1

codes and expand the coding to cover the larger number of characters. Since every endeavor was being made to implement standards, however, the Library decided to use the standard 7-bit USA Standard Code for Information Interchange (ASCII) issued by the United States of America Standards Institute (USASI). The proposed code configuration was based on an orderly and logical expansion and contraction of the 7-bit ASCII to 8-bit and 6-bit codes.

An 8-bit expanded ASCII (Figure 24) was the basic configuration. The 7-bit ASCII char-

Proposed Extended

ASCII

Character Set

—— Standard 6-bit set
-------- Non-standard set 1
·········· Non-standard set 2

8 7 6 5 BITS		0 0 0 0	0 0 0 1	0 0 1 0	0 0 1 1	0 1 0 0	0 1 0 1	0 1 1 0	0 1 1 1	1 0 0 0	1 0 0 1	1 0 1 0	1 0 1 1	1 1 0 0	1 1 0 1	1 1 1 0	1 1 1 1
		0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
0 0 0 0	0	NUL	DLE	SP	0	@	P	`[2]	p			ʼ	ʻ			ˀ	¸
0 0 0 1	1	SOH	DC1	!	1	A	Q	a	q			Ł	ł			ˋ	˛
0 0 1 0	2	STX	DC2	"	2	B	R	b	r			Ø	ø			ˊ	.
0 0 1 1	3	ETX	DC3	#	3	C	S	c	s			Đ	đ			^	..
0 1 0 0	4	EOT	DC4	$	4	D	T	d	t			Þ	þ			~	˳
0 1 0 1	5	ENQ	NAK	%	5	E	U	e	u			Æ	æ			—	=
0 1 1 0	6	ACK	SYN	&	6	F	V	f	v			Œ	œ			˘	_
0 1 1 1	7	BEL	ETB	'	7	G	W	g	w			ʹ	ʺ			•	̦
1 0 0 0	8	BS	CAN	(	8	H	X	h	x			•	ı			..	¿
1 0 0 1	9	HT	EM	)	9	I	Y	i	y			ƀ	£			ˇ	˘
1 0 1 0	A	LF	SUB	*	:	J	Z	j	z			®	ŏ			˚	⌢
1 0 1 1	B	VT	ESC	+	;	K	[	k	{[5,1]			±				ʿ	‿
1 1 0 0	C	FF	FS	,	<	L	\[1]	l	|[4,1]			Ơ	ơ			ʾ	
1 1 0 1	D	CR	GS[3]	—	=	M	]	m	}[5,1]			Ư	ư			ʼ	
1 1 1 0	E	SO	RS[3]	.	>	N	^[2]	n	~[2]							˝	ʼ
1 1 1 1	F	SI	US[3]	/	?	O	_[2]	o	DEL							̐	..

4 3 2 1
BITS

Key:

[1] Not on proposed print train.
[2] Redefined elsewhere in the set.
[3] To be used as terminators or delimiters.
[4] To be used as escape codes.
[5] To be used as shift codes for 6-bit set (non-locking) 7B shifts to non-standard set 1; 7D to non-standard set 2.

FIGURE 24.—*Extended USA Standard Code for Information Interchange*

acter set was left intact. Some of the standard characters such as the braces or the backwards slash were not used in the library character set. No characters, however, were substituted for these code positions. Other characters such as diacritical marks were left in their standard position (unused) and duplicated in another portion of the code set that was reserved for special characters and diacriticals.

The 7-bit code was derived from the 8-bit code by removing the 8th bit. Those columns which previously had a zero (0) in the 8th bit (columns 0-7) were considered the standard 7-bit set. Those with a 1 in the 8th bit (columns 8-F) were considered the nonstandard set. An SO (shift out) control character was used to go from the standard set to the nonstandard. The code stayed in the nonstandard mode until an SI (shift in) control character was reached.

The 6-bit code was derived by removing the 6th and the 8th bits. The 6th bit code consisted of the following sets:

Columns 2, 3, 6 & 7=Standard set
Columns 0, 1, 4 & 5=Nonstandard set (a)
Columns 8, 9, C & D=Nonstandard set—unused (b)
Columns A, B, E & F=Nonstandard set (c)

Three characters in the standard set—3C, 3D, and 3E were used as nonlocking precedence codes.

Subscript, superscript, and Greek characters were designed as alternate graphic sets in both the 6-bit and 8-bit codes. These alternate sets were reached by Escape sequences. These sequences are not yet defined, but for the purposes of description were called ESCA (for superscript), ESCB (for subscript) and ESCG (for Greek characters).

A proposed print train was designed based on the projected character set. In order to reduce the number of graphics, a number of compromises were made. For example, the cedilla (̧), and the left hook (,) will print as the same character even though they have different code representations on the tape. The proposed print train is shown in Table 11. In the final analysis, the actual design of the print train will be determined by what is feasible in terms of unique characters, cost for graphic artwork, and printing speed. Several studies are now underway at the Library to determine the effect of an expanded print train on printing speed.

The proposed character set and print train are being studied by the ALA Machine Readable Cataloging Format Committee.* It is hoped that by the time this report is published, the character set will be accepted by the library community as a standard.

TABLE 11.—*Proposed Library Print Train*

	CHARACTER	NAME		CHARACTER	NAME
1	‡*	Double dagger	22	6	
2	!	Exclamation point	23	7	
3	"	Quotation marks	24	8	
4	#	Number or sharp sign	25	9	
5	$	Dollar mark	26	:	Colon
6	&	Ampersand	27	;	Semicolon
7	'	Apostrophe	28	?	Question mark
8	(	Parenthesis	29	@	At sign
9	)	Parenthesis	30	A	
10	*	Asterisk	31	B	
11	+	Plus	32	C	
12	,	Comma	33	D	
13	-	Hyphen or minus	34	E	
14	.	Period	35	F	
15	/	Slash	36	G	
16	0		37	H	
17	1		38	I	
18	2		39	J	
19	3		40	K	
20	4		41	L	
21	5		42	M	
			43	N	
			44	O	

* A printing delimiter.

---

* The ALA Machine Readable Cataloging Format Committee was an ad hoc committee formed in 1967 by the Information Science and Automation Division, with representatives from the Reference Services Division and the Resources and Technical Services Division, to consider establishing library standards for machine-readable formats and character sets.

TABLE 11.—*Proposed Library Print Train*—Continued

	CHARACTER	NAME
45	P	
46	Q	
47	R	
48	S	
49	T	
50	U	
51	V	
52	W	
53	X	
54	Y	
55	Z	
56	[	Bracket
57	]	Bracket
58	a	
59	b	
60	c	
61	d	
62	e	
63	f	
64	g	
65	h	
66	i	
67	j	
68	k	
69	l	
70	m	
71	n	
72	o	
73	p	
74	q	
75	r	
76	s	
77	t	
78	u	
79	v	
80	w	
81	x	
82	y	
83	z	
84	ʼ	Alif
85	Ł	Polish L, uppercase
86	Ø	Scandinavian O with slash
87	Đ	D with cross bar, uppercase (eth)
88	Þ	Icelandic thorn, uppercase
89	Æ	
90	Œ	
91	ʹ	Miagkii znak
92	•	Dot in middle of line
93	♭	Musical flat
94	®	Subscript patent mark
95	±	Plus or minus
96	<	Less than
97	=	Equals
98	>	Greater than
99	ʻ	Ain
100	ł	Polish l, lowercase
101	ø	Scandinavian o with slash, lowercase
102	đ	D with cross bar, lowercase (eth)
103	þ	Icelandic thorn, lowercase
104	æ	
105	œ	
106	ʺ	Tverdyi znak
107	ı	Turkish i, lowercase
108	£	British pound
109	α	Alpha
110	β	Beta
111	γ	Gamma
112	ˋ	Grave
113	ˊ	Acute
114	^	Circumflex
115	~	Tilde
116	¯	Macron
117	˘	Breve
118	˙	Superior dot
119	¨	Umlaut or Dieresis
120	ˇ	Hacek
121	˚	Circle or Angstrom
122	︠	Ligature
123	︡	
124	ʼ	High comma
125	˝	Double acute
126	̐	Candrabindu
127	¸	Cedilla or left hook
128	˛	Right hook
129	.	Dot below character
130	..	Double dot below character
131	˳	Circle below character
132	‗	Double underscore
133	_	Underscore
134	$^{+}$	Superscript plus
135	$^{-}$	Superscript minus
136	$^{(}$	Superscript open parens
137	$^{)}$	Superscript closed parens
138	^{0}	Superscript numbers
139	^{1}	
140	^{2}	
141	^{3}	
142	^{4}	
143	^{5}	
144	^{5}	
145	^{7}	
146	^{7}	
147	^{9}	
148	$_{+}$	Subscript plus
149	$_{-}$	Subscript minus
150	$_{(}$	Subscript open parens
151	$_{)}$	Subscript closed parens
152	$_{0}$	Subscript numbers
153	$_{1}$	
154	$_{2}$	
155	$_{3}$	
156	$_{4}$	
157	$_{5}$	
158	$_{6}$	
159	$_{7}$	
160	$_{8}$	
161	$_{9}$	
162	Ơ	
163	Ư	
164	ơ	
165	ư	
166	ˀ	Pseudo question mark

Cost Models

Introduction

The need for determining what the processing rates and costs would be for an operational MARC system was evident during the early development stages of the pilot project. This information was necessary for planning purposes and could not be obtained with any confidence or precision by making estimates at that time. No comparable substantive data had been published in the literature or made available by other groups. Consequently, it was decided to monitor the operations of the MARC pilot and to develop a cost model of this system, identifying the unit costs associated with many of the suboperations of this system. This would permit informed decisions to be made regarding various subsystem alternatives. Changes made to the system could be monitored and evaluated in an operational environment. The result of a change would be reflected in a model and management would have a tool for analysis. In addition, extrapolations could be made for volumes and types of processing beyond those encountered in a pilot operation.

The original cost model for the period July 7, 1966 * to June 2, 1967 was prepared by Programming Services, Inc., and some of their conclusions are shown in the following pages. This model provided the framework for later cost analyses when the project became more stable. After June 1967, these analyses were prepared by members of the MARC staff. The results of these later cost analyses are also described.

After the decision was made to develop the cost model, arrangements were made to record the operations data from the MARC pilot. Data were recorded on control sheets attached to batches of material that were processed through various stages in the operation. An example of such a batch control sheet is shown in Figure 25. At each stage in the processing, personnel recorded their initials, the date, and the elapsed time for the completion of that part of the process. The sheets were collected at weekly intervals and the data punched into cards for subsequent computer processing and plotting. Recording the necessary information required nominal time on the part of the MARC personnel.

Because the MARC Pilot System was in a constant state of development, it was continuously being changed as new techniques and procedures were tried. Consequently, there was no single pilot system—there was, in fact, a continuously evolving system during the period of July 1966 to July 1967. The first seven months were devoted to training and experimentation. After January 30, 1967, the operation became more stable, and therefore only the costs of the second period from January 31 to June 2, 1967, are shown. The second period was divided into two phases for purposes of comparison:

Phase	*In Operation*
A	January 31, 1967, to March 6, 1967
B	March 7, 1967, to June 2, 1967

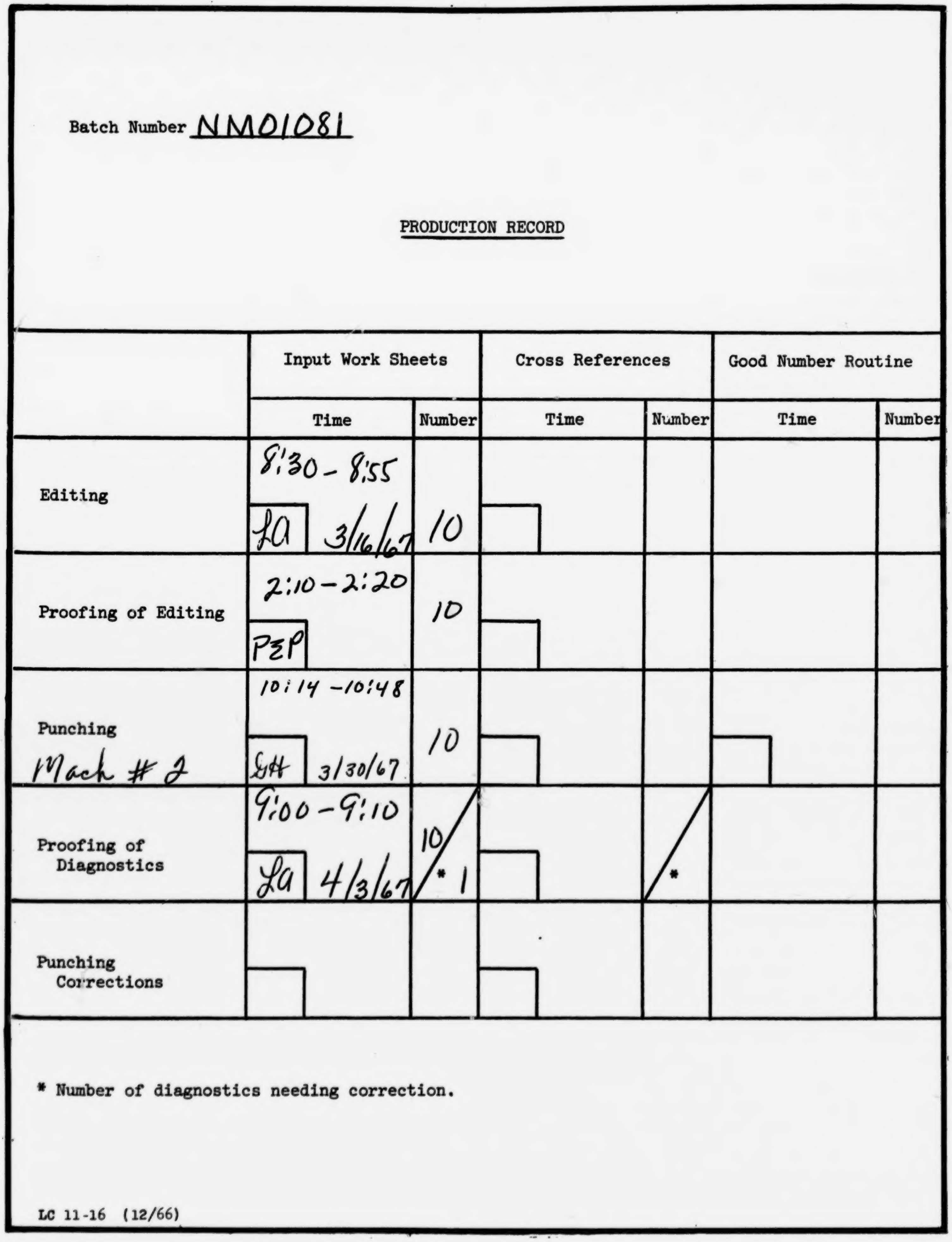
Analysis was restricted to the costs associated with production of the machine-readable cataloging record. The costs of the original cataloging, the costs for distributing multiple copies of the machine records, and the costs of producing the cross-reference records, were not included.

Basic Assumptions

It was assumed that the processing rates achieved during the period when batch times were recorded could not be sustained throughout the entire working day. There would be interruptions and other relief activities not completely represented in the recorded batch times. For that reason, an all-day effective sustained rate would in actual practice be something less

---

* The cost model was introduced in July 1966, four months before the actual implementation of the project.



Batch Number NM01081

PRODUCTION RECORD

	Input Work Sheets		Cross References		Good Number Routine	
	Time	Number	Time	Number	Time	Numbe
Editing	8:30 - 8:55 La 3/16/67	10				
Proofing of Editing	2:10 - 2:20 PEP	10				
Punching Mach # 2	10:14 - 10:48 GH 3/30/67	10				
Proofing of Diagnostics	9:00 - 9:10 La 4/3/67	10 / * 1		/ *		
Punching Corrections						

* Number of diagnostics needing correction.

LC 11-16 (12/66)

FIGURE 25. *Original Batch Control Sheet*

than the rates recorded in the batch mode. The analysis arbitrarily assumed that the all-day sustained rate would be 80 percent of the recorded batch rate. This correction is applied to the data at a later point in this report.

Staff time for system development and evaluation was not included in these estimates since this was a transient and peripheral activity not directly related to the production process and diminished to a relatively low level of effort after the system was fully operational. Training costs were also excluded wherever possible, but the data used in this analysis did include some work from staff members receiving on-the-job training in the production operation.

All work was done on a single shift except for computer processing.

Only costs that were directly identified with MARC production were included. General overhead costs such as rent, heat, light, employee benefits, and non-MARC labor were excluded from the analysis.

The computer equipment used at the Library of Congress for its regular data processing was made available to the MARC project. The rental cost for that configuration was $8,400 per month based on 176 hours per month of metered computer time. This amounted to a rate of $8,400/176=$47.73 per hour. MARC processing was charged at that rate for actual computer time used.

In addition to computer time, an operator was used to run the MARC programs. His salary level, GS-5 ($5,331/year) with a 10 percent differential for night work, amounted to an hourly rate of $5,864/2,000=$2.93 per hour. MARC processing was charged on an hourly basis for the operator time used in the actual processing of MARC tapes.

Salary levels were based on the then current Government Service (GS) pay schedules.

All workers were assumed to be fully employed at MARC-related tasks and working at the rate established. That is, there was no idle time other than that previously mentioned for minor interruptions.

Cost Model Results Based on Data Prepared By Programming Services, Inc.

What follows is a detailed description of the MARC input system giving the rates and costs experienced for each process. The data were separated into two periods of time: Phase A from January 31 to March 6, 1967, and Phase B from March 7 to June 2, 1967. The processes varied from system to system only as a result of some change in grades of personnel performing specified tasks and a general improvement in processing rates because of increased familiarity with the tasks.

1. Some supervisory and other indirect supporting activities were required to assist the regular MARC production effort. This included such things as staff supervision, development of new publisher and location codes, logging of input receipts and selection of items to receive MARC processing, preparation of batches and batch sheets, sorting of input records by LC card number, assignment of batch numbers, routing of work to the proper people, and filing worksheets after processing.

One person handled the material routing and record control functions. The control clerk used during the entire period was a GS-5 at an annual salary of $5,331. Control costs during Phase A, a 5-week period in which 2,180 titles received their initial coding, came to $0.235 a title. In the 13-week period covered by Phase B when 8,751 titles received their initial coding, control costs decreased to $0.152 per title.

Some editorial time was spent on the development of new publisher and location codes for the MARC records, resulting in the following data. Phase A: 73.84 manhours at the GS-4 to GS-11 salary levels were spent to develop new codes for 921 publishers at an average rate of 12.5 codes per hour. Approximately 2,180 titles received their initial coding at a total labor cost of $196.46 or a per title cost of $196.46/2,180=$0.09.* Phase B: 50.80 manhours at the GS-4 salary level were spent in developing new codes for 263 publishers, at an average rate of 5.2 codes per hour. Approximately 8,751 titles received their initial coding for a total labor cost of $121.41, a per-title cost of $121.41/8,751=$0.014.

Some editorial time was also spent on a variety of miscellaneous tasks such as checking place names in the Library's Geography and Map Division or reviewing special problems. This type of effort resulted in the following costs. Phase A: 6.23 hours of labor by people at the GS-4 to GS-11 levels produced a total labor cost of $24 over a period when 2,180 titles were receiving their initial coding. The result was a cost of $24/2,180=$0.011 per input title. Phase B: 29.42 hours of labor were reported. The work, done by people at the GS-4 to GS-11 salary levels, cost $96.51 over a period when 8,751 titles received their initial coding. This resulted in a prorated cost of $96.51/8,751=$0.011 per input title.

---

*During this period the backlog of titles needing new publisher codes was finally processed. This accounts for the large number of codes per title in Phase A.

General supervisory effort was provided by different people during the course of the project, and the amount was not recorded. For the purpose of this analysis, it was assumed that one full-time supervisor at the GS-11 rate ($9,221/year) was required for this purpose. Using this estimate, supervisory costs during Phase A, a 5-week period in which 2,180 titles received their initial coding, came to $0.407 a title. In the 13-week period covered by Phase B when 8,751 titles received their initial coding, supervisory costs amounted to $0.263 per title.

The supervision and various other indirect labor costs described above amounted to the following total per title:

	Phase A	Phase B
Control clerk	$0.235	$0.512
Establishment of publisher and location codes	0.090	0.014
Other editorial tasks	0.011	0.011
Supervision	0.407	0.263
Total per title	$0.743	$0.440

2. The input catalog records were obtained from the Library's Card Division by copying completed manuscript cards onto worksheet forms with a copying machine. In addition to running the machine, the operators, all at the GS-3 and GS-4 salary levels, also checked the quality of the output copy, checked the back of the card for additional information, and marked some appropriate blanks on the worksheet. Phase A showed a total of 4,071 titles copied in 28.43 hours, for an average copying rate of 143.19 titles per hour. The total recorded labor cost for these titles was $61.15 or an overall copying labor cost of $61.15/4,071=$0.015 per title. In Phase B, 11,985 titles were copied in 57.36 hours for an average rate of 208.94 titles per hour. The total recorded labor cost was $123.91 or an overall cost of $123.91/11,985=$0.10 per title.

3. The meter charge and cost of supplies for the copying equipment amounted to approximately $0.05 per page, including $0.035 for machine rental, $0.015 for paper, toner, drum replacement, etc. The Library of Congress Photoduplication Service actually charged MARC $0.06 per page for all copies made in the Library. Because continuation pages were necessary to show data recorded on the reverse side of manuscript cards, the copying process averaged more than one copied page per title.

	Total Titles	Total Copies	Copies Per Title
Phase A	4,071	4,383	1.08
Phase B	11,985	13,068	1.09

The above figures resulted in an average copying materials and equipment cost per input title of $0.06×1.08=$0.065.

4. The worksheets received preliminary coding from an editor (usually a junior editor), who reviewed them and assigned and transcribed publisher codes. The worksheets then went on to the professional staff, who completed the coding. On this process the following data were collected.

Phase	Salary Levels	Titles Coded	Hours Spent	Coding Rate	Total Cost	Unit Cost
A	GS-4 to GS-11	2,650	36.96	71.7 per hour	90.71	$0.034
B	GS-4 to GS-7	7,115	82.59	86.1	205.41	$0.029

5. After preliminary coding, the actual editing took place. The editors read the manuscript cards that had been copied on the worksheets and then recorded the appropriate tags, editorial corrections, and typing instructions. The following data were collected on this activity:

Phase	Salary Levels	Titles Edited	Hours Spent	Editing Rate	Total Cost	Unit Cost
A	GS-5 to GS-11	2,180	104.05	20.95 per hour	$329.78	$0.151
B	GS-5 to GS-11	8,751	324.32	29.68	$980.68	$0.112

6. Information on the worksheet was then typed on the paper-tape typewriter in order to obtain a punched paper tape. The typing was done in a highly formatted manner, a control tape entering some of the codes. The following amount of data was recorded on the typing, all of which was done at the GS-4 salary level.

Phase	Titles Typed	Hours Spent	Typing Rate	Total Cost	Unit Cost
A	2,478	232.63	10.65	$555.52	$0.224
B	9,371	709.82	13.20	$1,695.05	$0.181

The typewriters were rented for $150 per month each, including maintenance. Two typewriters were used during the project. During the five weeks of Phase A when 2,478 titles were typed, the cost of the two machines was $346.10 or $0.140 per title. In the 13 weeks of Phase B, 9,371 titles were typed at a machine cost of $899.86 or $0.096 per title. This machine cost included initial, correction, and verification typing.

7. The punched paper tapes were forwarded daily to the computer center for a number of separate processing operations including such things as transcription to magnetic tape, some validity checking, and a formatted printout. Data were collected to describe the machine time used for the daily MARC processing done during that time. Unfortunately these processing runs often performed many other tasks, e.g., processing cross-reference records, so that it did not seem possible to determine exactly how much of the recorded time should be assigned to the processing of new MARC records. For that reason, it was decided in this cost analysis to use data from runs that processed MARC records alone. A limited number of runs of this type were recorded during the period covered by each system. From them the following data, based on machine time used, were collected:

Phase A (data based on 11 machine runs):

$$\frac{\$47.73}{\text{hr.}} \times \frac{11.85 \text{ hours}}{1{,}529 \text{ titles}} = \$0.370 \text{ per input title}$$

Phase B (data based on 46 machine runs):

$$\frac{\$47.73}{\text{hr.}} \times \frac{64.03 \text{ hours}}{8{,}756 \text{ titles}} = \$0.349 \text{ per input title}$$

These figures should be used with some caution and *as an order of magnitude estimate only,* since they are the least precise of all the data used in this cost analysis for the following reasons: 1) The time recorded on the computer run sheets is not reliable because it is not always clear that the computer was actually being used for the entire period charged. Some of the time included production, testing, reruns, and setups. 2) The computer time was based on times recorded by the operators, rather than the time actually indicated on the time meter of the equipment. 3) There were some uncertainties in the numbers of new MARC records actually processed during each run. Some of the recorded numbers may be in error.

As mentioned earlier, the time of a computer operator was also charged to this processing. At an hourly rate of $2.93 this resulted in the following costs:

Phase	Titles Processed	Hours Spent	Unit Cost
A	1,529	21.80	$0.042
B	8,756	64.03	$0.021

8. A diagnostic listing was the first product after a catalog record had been entered into the computer system. These diagnostics were returned to the editors for proofreading, resulting in the data which follows:

Phase	Salary Levels	Titles Checked	Hours Spent	Proofreading Rate	Total Cost	Unit Cost
A	GS-5 to GS-11	1,819	54.46	33.40 per hour	$175.04	$0.096
B	GS-4 to GS-11	9,249	260.15	35.55	$801.35	$0.087

9. The diagnostic records that were found by the proofreader to be correct were endorsed for transfer from the interim tape to the master distribution tape. As they accumulated from the proofreaders, batches of these correct diagnostics were given to the paper tape typists who transcribed the LC card number on punched paper tape to make a list of correct MARC records. All of this work was done at the GS-4 level. In Phase A, 2,355 numbers were typed in eight hours for a typing rate of 294 numbers per hour, and proofread in 2.63 hours for a typing rate of 896 titles per hour. The unit cost totalled $0.011. In Phase B, 7,391 numbers were typed in 28.3 hours for a typing rate of 261 numbers per hour, and 5,340 numbers were proofread in 3.16 hours for a rate of 1,690 numbers proofread per hour. The unit cost was $0.01.

10. The diagnostic records that were found to contain errors were corrected by the proofreaders and then returned to the typists. The following data reflect the typing of necessary corrections; all work was done at the GS-4 level.

Phase	Titles Corrected	Hours Spent	Typing Rate	Total Cost	Titles Edited	Unit Cost
A	690	17.49	39.45 corrections	$41.77	2,180	$0.019
B	4,204	92.55	45.42 per hour	$221.01	8,751	$0.025

11. The paper tapes from the correction typing effort were then submitted for computer processing and the preparation of another diagnostic printout. These printouts were then proofread against the corrections on the old printouts. The correction proofreading was faster than the original proofreading, since fewer items of information on the record needed to be checked.

Phase	Salary Levels	Titles Proofread	Hours Spent	Proofreading Rate	Total Cost	Titles Edited	Unit Cost
A	GS–5 to GS–11	734	11.82	62.10 per hour	$39.81	2,180	$0.018
B	GS–5 to GS–11	3,790	57.12	66.35	$190.74	8,751	$0.022

12. The material costs for catalog record processing consisted primarily of paper tape and computer printing paper. Preliminary estimates indicated that an average of 500 characters were punched for each title, including the basic catalog record, corrections, subsequent correct record number, and associated tape leaders and other waste tape. Based on these figures, the paper tape cost was:

$$\frac{\$1.00}{1{,}000\text{ ft.}} \times \frac{\text{ft.}}{120\text{ char.}} \times \frac{500\text{ char.}}{\text{title}} = \$0.004\text{ per title}$$

The diagnostic printouts were made with one title per page. This resulted in a printing paper cost of:

$$\frac{\$2.22}{1{,}000\text{ pages}} \times \frac{1\text{ page}}{\text{title}} = \$0.002\text{ per title}$$

The total materials cost, applied consistently to each of the two discrete systems came to $0.006 per title.

13. The total input processing cost to put the original catalog records on magnetic tape thus consisted of the elements described above with a correction based on the assumption that the all-day sustained labor rate would be 80 percent of the recorded batch rate (see Table 12). Any moderate increase in the volume of similar material to be processed by each of the systems described should have the following incremental cost per title:

Phase A: $2.260 per title
Phase B: $1.687 per title

TABLE 12.—*Analysis of Input Processing Costs*

	Cost per Input Title Phase A (31 Jan 1967–6 Mar 1967)	 Phase B (7 Mar 1967–2 Jun 1967)
Labor Cost Elements (Net)		
Supervision & indirect support	$0.743	$0.440
Manuscript card copying	.015	.010
Assign publisher codes to worksheets	.034	.029
Initial worksheet editing	.151	.112
Initial typing	.224	.181
Computer operator	.042	.021
Diagnostic proofreading	.096	.087
Good number typing and proofreading	.011	.010
Correction typing	.019	.025
Correction diagnostic proofreading	.018	.022
	$1.343	$0.937
Total Labor Cost (Effective) (allowing for only 80% productive time)	$1.343/0.8 =1.679	$0.937/0.8 =1.171
Materials and Other Direct Cost Elements		
Dura machine rental	$0.140	$0.096
Copying machine and supplies	.065	.065
Computer processing	.370	.349
Expendable materials	.006	.006
	$0.581	$0.516
Total Labor and Other Direct Costs:	$2.260	$1.687

Cost Models Prepared by the MARC Staff

The performance data recorded in this report indicate a gradual improvement in the performance of several subprocesses such as editing and typing. The cost figures in Table 12 also reflect this improvement. One limitation of this report, however, is its dependence on the system described. During the MARC Pilot Project no better information was available on which to base rational judgments about the nature of successively improved methods of inputting catalog records in digital form. But by June 1967, when the original cost model was completed, the MARC staff had a full year of experience with the processes and, as a result, two major problems were recognized: first, that overall costs could not drop further without additional modifications of the processing procedures or equipment and second, that the method of analyzing costs needed improvement to eliminate those areas where costs could not be measured with complete accuracy.

The first problem was answered by a thorough analysis of the work procedures of each employee and a careful evaluation of the optimum GS grade necessary for each task consistent with acceptable work performance. The demand for a high level of accuracy resulted in reinstatement of the proofreading after editing. The initial coding was assigned to the editors, and the control clerk, who had formerly performed that process, was gradually relieved of various tasks which were considered nonessential. By December 1967 the production process was operating smoothly without a control clerk, and the supervision required had dropped considerably from the level recorded in this report. Furthermore, changes in the programs had resulted in a significant drop in the computer time necessary to process MARC records, which in turn resulted in a reduction in the cost of that part of the processing operation. It is now once again true that further reductions in cost are not possible without modification of the processing system, and plans are already being made to accomplish this. It is hoped that in the near future a redesigned form for collecting catalog data in the Library of Congress will make it possible to delegate part of the MARC data preparation process to the Library's professional catalogers.

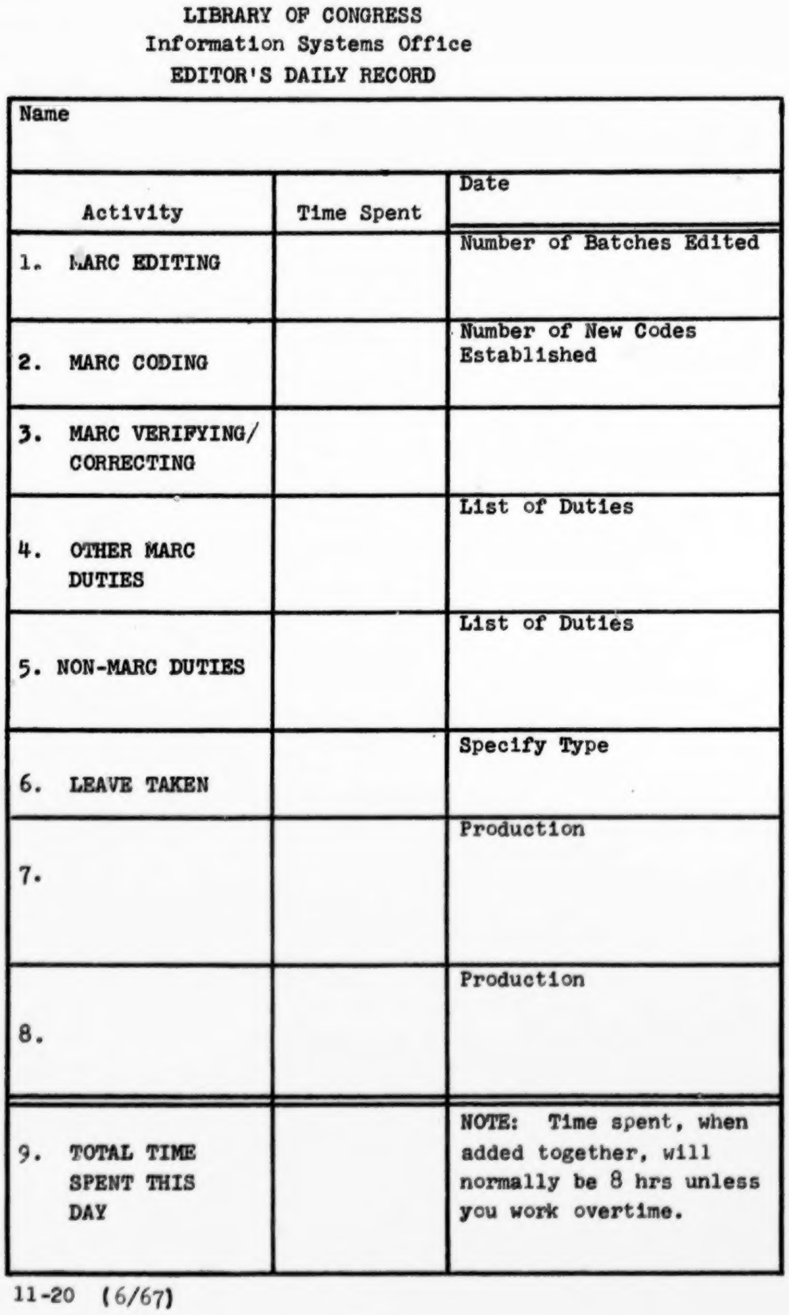
Improvement of the method of analyzing costs was the second major task. That was accomplished by introducing a series of new forms (see Figures 26 and 27) on which to collect actual time and production reports from every worker every day. All time spent on the job, including nonproductive time, was included in the accounting, as was the amount of work produced during that day. These records corrected several weak features in the original cost model and provided actual production costs over a series of one-month periods. As a result, supervision costs could be assigned with far greater accuracy, it was no longer necessary to estimate the number of records produced by each paper-tape typewriter during each month, computer time could be measured more accurately and separated from computer operator time, the effects of eliminating the control clerk and making other modifications could be accurately monitored and measured within the operational framework, and the arbitrary reduction of maximum production levels by 80 percent could be abandoned in favor of the measurement of actual production rates and costs. The basic structure of the cost models was retained, having proved a useful and expedient method of subdividing the MARC production process, and costs were, as before, assigned to each function and subprocess in the system.

As a result of the changes in the work procedures and the revised method of cost analysis, new sets of cost figures were obtained on a monthly basis. The month of October 1967 provides an illustration of how this new system operated. During that month 2,901 records were processed, revealing the following cost elements:

Supervision and indirect support	$0.291
Manuscript card copying	.007
Assign publisher and place codes	.042
Initial worksheet editing and proofreading	.162
Initial paper tape punching	.244
Computer operator	.041
Diagnostic proofreading	.125
Verified numbers and correction punching	.039
Total labor cost	$0.951
Paper tape typewriter rental	$0.156
Copying machine and supplies	.065
Computer processing	.387
Paper materials	.015
Other direct costs	$0.623
Total labor and other direct costs	$1.574

Since these figures were based on actual performance during the month, no de-rating or other adjustment was necessary.

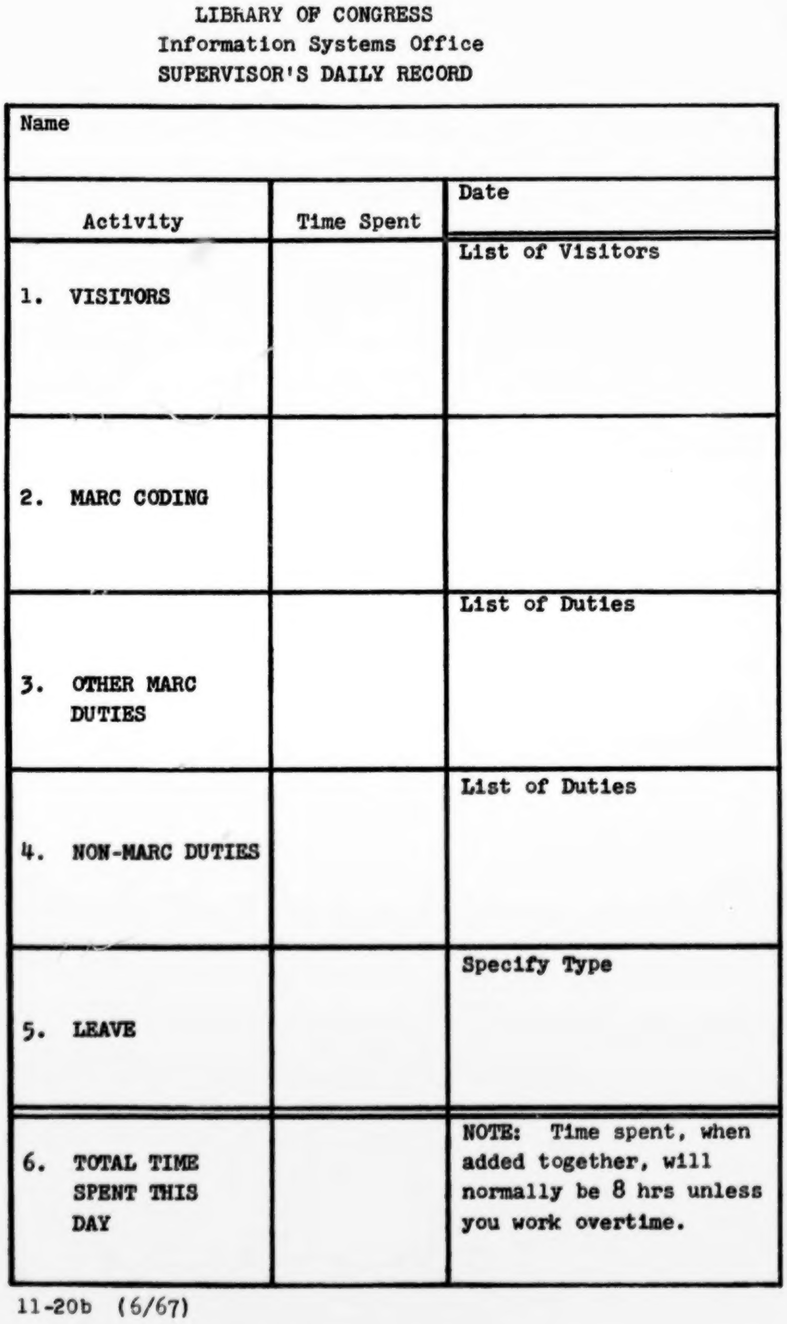
The above figures revealed that, contrary to earlier belief, the cost of supervision was not hugely disproportionate, but that paper tape punching, computer processing, and supervi-



LIBRARY OF CONGRESS
Information Systems Office
EDITOR'S DAILY RECORD

Name		
Activity	Time Spent	Date
1. MARC EDITING		Number of Batches Edited
2. MARC CODING		Number of New Codes Established
3. MARC VERIFYING/ CORRECTING		
4. OTHER MARC DUTIES		List of Duties
5. NON-MARC DUTIES		List of Duties
6. LEAVE TAKEN		Specify Type
7.		Production
8.		Production
9. TOTAL TIME SPENT THIS DAY		NOTE: Time spent, when added together, will normally be 8 hrs unless you work overtime.

11-20 (6/67)



LIBRARY OF CONGRESS
Information Systems Office
SUPERVISOR'S DAILY RECORD

Name		
Activity	Time Spent	Date
1. VISITORS		List of Visitors
2. MARC CODING		
3. OTHER MARC DUTIES		List of Duties
4. NON-MARC DUTIES		List of Duties
5. LEAVE		Specify Type
6. TOTAL TIME SPENT THIS DAY		NOTE: Time spent, when added together, will normally be 8 hrs unless you work overtime.

11-20b (6/67)

FIGURE 26.—*Time and Production Reports for MARC Interim System*

LIBRARY OF CONGRESS
Information Systems Office
PAPER TAPE TYPIST'S DAILY RECORD

Name		
Activity	Time Spent	Date
1. NEW MARCs		Number of Batches of NEW Records Punched
2. MARC CORRECTIONS/ VERIFICATIONS/ DELETIONS		
3. OTHER DUTIES		List of Duties
4. LEAVE TAKEN		Specify Type
5.		Number of Batches of NEW Records Punched
6.		
7. TOTAL TIME SPENT THIS DAY		NOTE: Time spent, when added together, will normally be 8 hrs unless you work overtime.

11-20a (6/67)

LIBRARY OF CONGRESS
Information Systems Office
COMPUTER OPERATOR'S DAILY RECORD

Name		
Activity		Date
1. DAILY MARCs		NEW Records Processed
METER Start	TIME Start	CORRECTED Records Processed
METER Stop	TIME Stop	VERIFIED Records Processed
2. WEEKLY MARCs		TOTAL Records Processed
METER Start	TIME Start	
METER Stop	TIME Stop	
3.		NEW Records Processed
METER Start	TIME Start	CORRECTED Records Processed
METER Stop	TIME Stop	VERIFIED Records Processed
4.		TOTAL Records Processed
METER Start	TIME Start	
METER Stop	TIME Stop	

11-20c (6/67)

FIGURE 27.—*Time and Production Reports for MARC Interim System*

sion were important cost elements which together contributed more than half the total cost of processing each MARC record. These elements were made the focus of attempts to improve the MARC input system, and changes in the control and batching procedures, the assignment of tasks to individual staff members, and the computer programs all contributed to increased efficiency and lower costs despite the 4-percent salary increase resulting from passage of the Federal Salary Act of 1967.

By the spring of 1968 it had become apparent that many of the input costs had achieved a high degree of stability. The major exceptions were the following:

1. Supervision costs had declined as production had risen. By March 1968 the MARC System Production Group was producing approximately 240 records each day under a single supervisor, whose pay at grade GS–12 ($11,461 yr.) resulted in a supervisory cost of $0.183 per record. Moreover, it seemed likely that this cost would drop even further if production could increase without requiring an additional supervisor, since the establishment of new codes had been discontinued and the job of the control clerk had been absorbed into the editors' functions.

2. The assignment of publisher codes had been discontinued in anticipation of the MARC II format, which does not include a code for publisher, and in response to the MARC participants, who had not found the code useful. The place code remained, but the function of assigning this code to each record was absorbed into the initial editing of the worksheet without apparent detriment to efficiency.

3. The cost of punching a MARC record, including all necessary corrections and verification, was reduced from $0.283 in October 1967 to $0.240 in March 1968. It was not expected that this cost could be lowered further.

4. The cost of computer processing was still an unreliable figure at best. Computer processing times are related to the size of the file being processed, and processing times range from seconds when the file is small to minutes as the file grows. Thus computer processing costs were large in February 1968 when volume 4 of the MARC data base neared its maximum size of almost 11,000 records; by March costs had dropped significantly solely because volume 5 had been started and the data base consisted of only a few thousand records. During March the cost of processing each MARC record averaged $0.318 including the entire procedure of sorting, printing, and processing each day's input and updating the master data base. Such inconsistencies will be avoided and processing times will drop when the MARC Distribution Service, which will deal only with noncumulated tapes, replaces the present pilot distribution service.

The cost of processing each MARC bibliographic record input in March was approximately as follows:

Supervision	$0.183
Manuscript card copying	.007
Initial worksheet editing and proofreading	.162
Initial paper tape punching	.207
Computer operator	.041
Diagnostic proofreading	.125
Verified numbers and correction punching	.033
Total labor cost	$0.758
Paper-tape typewriter rental	$0.156
Copying machine and supplies	.065
Computer processing	.318
Paper materials	.015
Other direct costs	$0.554
Total labor and other direct costs	$1.312

The gradual decrease in total MARC input cost from $2.260 (in March 1967) to $1.312 (in March 1968) is, in a large measure, attributable to these cost models. The cost models described in this report and the cost analyses made since June 1967 have provided an increasingly reliable foundation for predicting future system costs and behavior. Furthermore, they have resulted in a production operation whose efficiency has gradually and steadily improved, bringing about a significant increase in the weekly production and distribution of MARC records and permitting the accurate prediction of staff, equipment, and budgetary needs for the Library of Congress' MARC Distribution Service.

Evaluation of MARC I and Comparison with MARC II

The MARC I format was shaped by many constraints. It was designed for testing and evaluation in an experimental project. Time limitations restricted the form of material to monographic and also dictated that the format be stabilized by April 1966 in order to begin programming. The format was structured so as to lend itself readily to ease in programming on an IBM System/360 Model 30 with 16K core memory. Finally, its design was influenced by the emphasis of the project on distribution of LC cataloging information.

The MARC Pilot System demonstrated the need for standardization in format design, character sets, codes, etc., for bibliographic description. The ever-increasing interest in MARC directed the thinking in the Library of Congress toward a bibliographic service that was broader in concept than a distribution service from the Library of Congress to user libraries.

The need for the capability of interchange of bibliographic information in digital form was being discussed in many circles. Therefore, the design of the MARC II placed emphasis on a format that would be an efficient means of communicating bibliographic description of all forms of material, i.e., monographs, serials, maps, music, etc. In concept, the format is based on three elements: structure, content, and coding. The structure provides the basic machine framework of the record. The content relates to the type of material in the record, the tags describing this material, and the coding is the binary configuration of the characters. The basic aim of MARC II format was to provide a structure which could be used with all types of material. This structure has been described in detail in the MARC II format report[6] issued by the Library of Congress.

Ideally, if the library community knew the data elements for all material that might be transmitted in digital form, the universe of such material could be studied at one time, and the content of the record could be standardized as well. Decisions could be reached on a generalized tagging scheme. This is not so, however, and we cannot wait for the ultimate. Therefore, it was assumed that each form of material would be studied by the people knowing that material the best, who would define the data elements, all within the same basic structure.

On the basis of the MARC pilot experience, the Library of Congress proceeded to define data elements for monographs. The intent was to design a format as "rich" as possible within the limitations imposed by the ability of the LC catalogers to provide the information and maintain the cataloging effort at the Library. This would enable the user to include or exclude data elements as he saw fit.

The design of the new format for monographs was considerably influenced by the evaluation of the MARC I format. This evaluation was supplied by the participants, by the LC staff, and by many other librarians and computer specialists who expressed interest in the project.

A summary of the more significant points brought out in the evaluation of the MARC I format follows:

Fixed Fields

1. Type of Main Entry. Experience at LC showed that dividing corporate names by type was more difficult than had been anticipated. In addition, the majority of the participants felt this fine breakdown was unnecessary. Consequently, MARC II will distinguish only personal names, corporate names (and as a subdivision of that, conferences) and uniform title headings. This information will be part of the tagging structure.

2. Form of Work. This type of information was expanded to cover other types of material and will be included in the legend in MARC II.

3. Bibliography Indicator. There was agreement that the form of content should be described more fully. Bibliographies will be one of a set of codes designated to describe the content of a work, e.g., catalogs, indexes, abstracts, etc.

4. Illustration Indicator. The need was felt for a more complete description of illustrations. Therefore, in MARC II, up to four types will be coded.

5. Language. There was general agreement that although two language fields were enough for the majority of works, it was desirable to have the ability to describe more languages if necessary. In MARC II a fixed field describes books in one language. When more than one language is described a variable field is used.

6. Place of Publication. It was felt that a coded place of publication was desirable but that it could be limited to country only.

7. Publisher. This code proved to be difficult to maintain and will be omitted from MARC II.

8. Height of Volume. This information will be specially identified in the variable fields by a delimiter and therefore will be omitted from the fixed fields.

Variable Fields

1. LC Call Number. A number of participants indicated that they would like to be able to distinguish between the class number and the book number. In the majority of the cases, the first alpha after the numeric portion of the class number begins the book number. In some cases, however, this alpha is a subclass of the class number. For this reason a delimiter will be inserted after the class number in MARC II.

2. Main Entry. The tag in MARC I indicated only that the field was a main entry. In MARC II the tags themselves will reflect the type of main entry. In addition, information regarding form of name will be shown in an indicator associated with the field.

3. Conventional or Filing Title. There was some confusion about this field since the MARC I record included both the conventional titles printed on the card and the filing titles that LC used in its own system but that did not appear on the LC printed card. In MARC II an indicator associated with the field will show which titles are of the type printed on the LC card and which reflect internal LC filing practice.

4. Imprint Statement. In MARC I, the imprint was divided into three subfields containing place, publisher, and date. Other imprint patterns such as place-date, place-publisher, place-publisher-date, and place-place-publisher-date, made sub-division into three fields difficult. With the advent of Title II–C, it is impossible to predict what patterns may appear in the imprint. In order to recognize the components of the imprint, MARC II will identify each unique component.

5. Collation Statement. The majority of the participants felt that the collation statement should be separated into its component subfields—pagination, illustrations, and size.

6. Notes. It became apparent that the identification of different types of notes would be useful. Therefore, MARC II will tag separately bibliography notes, content notes, etc.

7. Subject Tracings. Because of time limitations no attempt was made to differentiate between kinds of subject tracings. In MARC II, a system of tags and coded delimiters will identify types of subject headings, e.g., personal name, geographic name, and their subdivisions, e.g., period subdivisions.

8. Copy Statement. In MARC II the LC call number will be separated from the rest of the copy statement by a delimiter.

The following is a gross comparison of the MARC I and MARC II formats for monographs.

MARC II

Structure

Field 1—The leader area is fixed and contains record length and a data description of the record that follows.

Field 2—A variable length directory composed of fixed length entries containing tag, length of field, and starting character position (SCP) of field. The last entry in the directory is followed by a field terminator.

e.g. 100 0406 00945
Tag Length SCP

MARC I

Structure

Field 1—Fixed fields contained block length, record length, control number (LC catalog card number), codes, and other fixed field indicators

Field 2 to n—Variable fields containing length of field, tag of field, and data which could be subdivided by delimiters.

MARC II	MARC I
Structure	*Structure*
Field 3 to n—Variable fields containing indicators and data which may be subdivided by delimiters and subfield codes. Each variable field is followed by a field terminator. Variable fields 001–008 have been designated as control fields. The fixed field bibliographic information recorded in the fixed fields in MARC I is recorded in a variable fixed field in MARC II.* Control fields do not have indicators.	
Tags	*Tags*
MARC II employs a three-character numeric tag. Additional information pertaining to a field is described in two characters (indicators) preceding the data in each variable field.	MARC I used a two-character numeric tag with a third character (alphanumeric) reserved to make explicit certain information about a field. When the third character did not contain explicit information, the character was coded as a blank.
Delimiters	*Delimiters*
Each subfield is identified by a code made up of a delimiter followed by an alphabetic character. This constitutes a subfield code.	Single delimiters were used to separate subfields.
Content	*Content*
Bibliographic Fixed Fields	*Bibliographic Fixed Fields*
Type of publication date	Type of publication date
Date 1	Date 1
Date 2	Date 2
Country of publication code (country only)	Place of publication code (country and city)
Illustration codes (up to four) choice made by ranking illustrations in priority order.	Illustration indicator (single indicator regardless of type of illustration except maps)
(Included in illustration code)	Map indicator
Intellectual level code	Juvenile indicator
Form of microreproduction code	---------------------
Form of contents code (bibliographies, catalogs, indexes, abstracts, dictionaries, etc.)	Bibliography indicator only
Government publication indicator	---------------------
Conference or meeting indicator	Conference or meeting indicator
Festschrift indicator	---------------------
Index indicator	---------------------
Main entry in body of entry indicator	---------------------
Fiction indicator	---------------------
Biography indicator	---------------------
---------------------	Publisher
(This information is made explicit in the tag for heading)	Type of main entry
(This information is made explicit in the tag for heading)	Form of work
Language code (if one language only; otherwise recorded in variable field)	Language indicator
	Language 1
	Language 2
(This information is made explicit in the tag for secondary entries)	Types of secondary entries
---------------------	Series indicator

---

*Fixed fields in MARC II are described as variable when defined for all types of records but are fixed in length when pertaining to a specific type of record.

MARC II	MARC I
Variable Fields	*Variable Fields*
Control Fields	
0 0 1 Control Number (LC catalog card number)	Recorded in fixed field
0 0 2 Sub-Record directory	---------------------
0 0 4 Cataloging source	---------------------
0 0 8 Fixed fields	Not variable in MARC I
0 0 9 Languages	Fixed field in MARC I
Control Numbers	
0 1 0 LC card number (for users who enter their own control number in field 001)	94 LC card number (only for suffix to LC card number)
0 1 1 Linking LC card number	---------------------
0 1 5 National bibliography number	83 National bibliography number
0 1 6 Linking NBN	---------------------
0 2 0 Standard book number	---------------------
0 2 1 Linking SBN	---------------------
0 2 5 Overseas acquisitions number (PL480, LACAP, etc.)	Bibliographic note (Tag 60)
0 2 6 Linking OAN number	---------------------
0 3 5 Local system number	---------------------
0 3 6 Linking local number	---------------------
0 3 9 Search code	---------------------
Knowledge Numbers	
0 5 0 LC call number (delimited between class number and book number. Indicator used to discriminate as to whether book is or is not in LC).	90 LC call number (no use of delimiter or indicator. NLC added to number when book not in LC).
0 5 1 Copy statement (delimited between LC call number and rest of field).	80 Copy statement (no use of delimiter)
0 6 0 NLM call number	---------------------
0 7 0 NAL call number	---------------------
0 7 1 NAL subject category number	---------------------
0 8 0 UDC number	---------------------
0 8 1 BNB classification number	---------------------
0 8 2 Dewey Decimal number	92 Dewey Decimal number
0 9 0 Local call number	---------------------
Main Entry	
1 0 0 Personal name	10 Main entry (type of main entry indicated in fixed field).
1 1 0 Corporate name	
1 1 1 Conference or meeting	
1 3 0 Uniform title heading	
Supplied Titles	
2 4 0 Uniform title (indicator used to show when printed)	15 Conventional or filing title (no use of indicator)
2 4 1 Romanized title	---------------------
2 4 2 Translated title	---------------------
2 4 3 Uniform title (collective works, reserved for British MARC)	---------------------
Title Paragraph	
2 4 5 Title	20 Title statement
2 5 0 Edition statement	25 Edition statement
2 6 0 Imprint (use of subfield codes allows any pattern to be recognized).	30 Imprint (could not recognize all possible patterns of place, publisher, date)
Collation	
3 0 0 Collation (delimited between pagination or volumes, illustration statement, height, and thickness).	40 Collation (no use of delimiters)

Collation—Continued

MARC II	MARC I
Variable Fields	*Variable Fields*
3 5 0 Bibliographic price	Part of imprint field (Tag 30)
3 6 0 Converted price	---------------------
Series Note	
4 0 0 Personal name—title (traced same)	50 Series note-Traced same (all series traced the same were identified by the same tag—further divided by use of 3d character in the tag to indicate author/title or title).
4 1 0 Corporate name—title (traced same)	
4 1 1 Conference—title (traced same)	
4 4 0 Title (traced same)	
4 9 0 Series untraced or traced differently	51 Series note-untraced or traced differently
Bibliographic Notes	
5 0 0 General notes	60 Notes (all notes were identified by the same tag. There was no differentiation between types of notes)
5 0 1 "Bound with" note	
5 0 2 Dissertation note	
5 0 3 Bibliographic history note	
5 0 4 Bibliography note	
5 0 5 Contents note (formatted)	
5 0 6 "Limited use" note	
5 2 0 Abstract	
Subject Added Entry	
6 0 0 Personal name	70 Subject tracing (all subject added entries were identified by the same tag. There was no differentiation between types of subject added entries)
6 1 0 Corporate name (excluding political jurisdiction alone)	
6 1 1 Conference or meeting	
6 3 0 Uniform title heading	
LC Subject Headings	
6 5 0 Topical	
6 5 1 Geographic names	
6 5 2 Political jurisdiction alone or with subject subdivisions	
Other Subject Headings	
6 6 0 NLM subject headings (MESH)	---------------------
6 7 0 NAL subject headings (Agricultural Biological Vocabulary)	---------------------
6 9 0 Local subject heading systems	---------------------
Other Added Entries	
7 0 0 Personal name	71 Personal name
7 1 0 Corporate name	72 Corporate name (further subdivided by use of 3d character in the tag to indicate Government body, society or institution, religious society or institution, miscellaneous corporate body)
7 1 1 Conference or meeting	
7 3 0 Uniform title heading	73 Uniform tracing
7 4 0 Title traced differently (if traced the same, shown by an indicator in title statement field)	74 Title tracing (if traced the same, field contained only a T; if traced differently contained a T and data)
7 5 0 Proper names not capable of authorship	---------------------
Series Added Entries	
8 0 0 Personal name—title (traced differently)	75 Series tracing (if traced the same contained an A or a T; if traced differently contained an A or a T and data. T indicated a title series A indicated an author/title series.)
8 1 0 Corporate name—title (traced differently)	
8 1 1 Conference or meeting—title (traced differently)	
8 4 0 Title (traced differently)	
9 0 0 BLOCK OF 100 NUMBERS FOR LOCAL USE	LOCAL USE NUMBERS (in fixed field)

In the design of MARC II, the tagging structure was based on assigning the primary digit a functional meaning and the secondary digits a "type of name" meaning. For example, the tag for a corporate name subject entry is 610. The first digit (in this case "6") describes the function of the name, i.e., it serves as a subject entry. The second digit (in this case "1") indicates that the name is a corporate name. This structure is true for all fields which are capable of being a heading in a file.

It should be emphasized that MARC II is a format designed as a vehicle for the transmission of bibliographic data. For retrieval purposes, the high-use data elements will probably be selected from the record and indexes constructed for searching.

Therefore, regardless of the tagging scheme employed, certain data elements will have to be selected from the record for retrieval purposes and a link to the master record provided. The important consideration is to ensure that the type of name can be easily recognized. Table 13 demonstrates that the ability is provided to recognize the function or the type of name.

TABLE 13.—*MARC II Tags*

Digit 1 = Function Digit 2 = Type of name Digit 3 is reserved for breakdown of type of name	Field	Digit 1 = Function Digit 2 = Type of name Digit 3 is reserved for breakdown of type of name	Field
BY TYPE OF NAME		BY FUNCTION	
100	Personal name main entry	100	Main entry personal name
600	Personal name subject entry	110	Main entry corporate name
700	Personal name added entry	130	Main entry uniform heading
400	Personal name series note		
800	Personal name series tracing	245	Title
110	Corporate name main entry	400	Series note personal name
610	Corporate name subject entry	410	Series note corporate name
710	Corporte name added entry	440	Series note title
410	Corporate name series note		
810	Corporate name series tracing	600	Subject personal name
		610	Subject corporate name
130	Uniform title main entry	630	Subject uniform heading
630	Uniform title subject entry		
730	Uniform title added entry	700	Added entry personal name
		710	Added entry corporate name
245	Title	730	Added entry uniform heading
740	Title added entry	740	Added entry title
440	Title series note		
840	Title series tracing	800	Series tracing personal name
		810	Series tracing corporate name
		840	Series tracing title

Potential Uses of Machine-Readable Data

The MARC Pilot Project was planned as an experiment to test the feasibility and utility of centrally producing machine-readable cataloging records and distributing these records to user libraries. The library community responded to MARC in terms requiring almost immediate modifications of the project definition, its life, and its implications. It became evident that what had been conceived as a small scale equipment would become the basis for a production operation on a much larger scale.

This enthusiastic response also made the realization of the need for standards more acute. The Library of Congress, with the cooperation of many libraries in this country and abroad, has developed these standards, and they are being adopted. Cataloging data, in machine-readable form, will become a national resource available to all. Standardized recording allows these data to be used with a minimum of redundancy in system design and programming efforts. Thus, large institutions and the smaller academic libraries, public libraries, and special libraries can use the information economically.

Standardization will have a significant impact on hardware costs. Agreement on a set of characters to express bibliographic data should result in the design of input and output devices to satisfy these requirements. The cost of development on the part of the manufacturer will benefit him and consequently these benefits should be felt throughout the library community. Standards in the representation of bibliographic description are a fundamental requirement for creating the environment in which these descriptions can be shared.

We are living in an era experiencing tremendous acceleration in many fields. Ours is a society that demands rapid access to large bodies of information. How can we proceed to satisfy the needs of the student, the scholar, the researcher during the period in which both intellectual and technical difficulties prohibit the development of a utopian library?

Efficiently designed and implemented technical processing centered about a computer-based data bank should make possible the fuller realization of goals long sought by libraries. Among them are the following:

1. Release of personnel to perform more intellectual tasks. Jobs now neglected for lack of personnel could be accomplished.
2. Economic utilization of a single record which could be manipulated to produce a variety of products, e.g., catalog cards, book catalogs, acquisitions forms, spine labels.
3. Elimination of the need for large manual files.
4. Up-to-date lists of current publications from which books could be ordered using a minimum amount of human effort.
5. Facilitation of the production of union catalogs.
6. The ability to select and modify data in a record to suit local user requirements.
7. Access to national bibliographic resources.

Title III of the Library Services and Construction Act as amended in 1966 provides Federal assistance to establish and maintain local, regional, State, or interstate cooperative networks of libraries for the systematic and effective coordination of the resources of school, public, academic, and special libraries and special information centers. The MARC Distribution Service makes the concept of cooperation more effective by the use of automated techniques. In addition, cooperative efforts provide the possibility of sharing computer facilities which might be too costly for small and medium-sized libraries.

The concept of sharing information via networks preceded MARC, but MARC has accelerated the planning and implementation. The cooperative efforts of the national libraries under the U.S. National Libraries Task Force on Automation and Other Cooperative Services indicate the awareness of the potential use of the new communications media. The New England Library Information Network (NELINET)

Project under the direction of the New England Board of Higher Education has developed the initial phases of a computer-based technical processing center. The source data used is MARC. Six New England university libraries are participating.

The Ohio College Library Center is in the planning stages for a technical processing center to serve 50 to 60 libraries. An extensive network for the libraries of the State of Washington has been proposed. Many other proposals, papers, and studies are under consideration.

The above description is not an attempt to provide a review of the current status of information networks nor pass over lightly the obstacles that still lie in our path. It is too easy to gloss over critical problems involved in the automation of libraries in our desire to approach our goals. The ability to separate fact from fancy is a necessary prerequisite for moving ahead.

Bibliography

1. Bush, Vannevar. As we may think. The Atlantic monthly, v. 176, July 1945: 101–108.
2. Buckland, Lawrence F. The recording of Library of Congress bibliographical data in machine form. Washington, Council on Library Resources, 1965. 54 p.
3. Avram, Henriette D., Ruth S. Freitag, *and* Kay D. Guiles. A proposed format for a standardized machine-readable catalog record; a preliminary draft. [Washington] Library of Congress, Office of the Information Systems Specialist, 1965. 110, 4, 22 p. (ISS planning memorandum no. 3)
4. Leach, Theodore E. A compendium of the MARC System. Library resources and technical services, v. 12, summer 1968: 250–275.
5. Rather, Lucia J. Special characters and diacritical marks used in roman alphabets. Library resources and technical services, v. 12, summer 1968: 285–295.
6. Avram, Henriette D., John F. Knapp, *and* Lucia J. Rather. The MARC II format; a communications format for bibliographic data. Washington, Library of Congress, 1968. 167 p.

Appendixes

Appendix A REPORTS OF PARTICIPANTS

Introduction

The section that follows contains the evaluations of the MARC Pilot Project written by the participants themselves. The only editing was stylistic in nature. All comments, including negative ones, concerning the project have been left intact. As stated by the Librarian of Congress in his forward to this report, "The Library of Congress, in its endeavor to serve the library community, therefore, feels that it is vital to progress to report all results, both good and bad, to serve as guides for future development of library automation. We need to know what went wrong as well as what went right with a process or technique in order to avoid repeating the mistakes of our predecessors."

The MARC Pilot Project submitted proposed guidelines to the participants to assist them in their evaluation. These guidelines precede the reports themselves in the following pages. Some of the participants followed the guidelines, others reported their findings in individual formats.

The principal criticisms of the project centered on the computer programs submitted to the participants for use at their local institutions and the quality of the data included in the MARC tapes.

No attempt is being made to argue that the programs supplied to the participants were efficient and well designed. They were not. It is only fair to the contractor, however, and in turn to the Library of Congress, to make the following statements about the programs. It would serve no useful purpose to turn back time to investigate how the chart describing the minimum computer configuration common to all participants was derived. The chart did reflect a 1401 8K memory with no special features. Therefore, this was the configuration used by the contractor as the base equipment for one set of software. What is not known is whether a participant made an error in reporting the information or the contractor made an error in recording the operation. The contractor did experience some on-the-job training, i.e., several of the programmers assigned to the project had never coded for a System/360. This situation was not uncommon early in 1966, and still exists today. It has confronted most people responsible for the design and implementation of software.

In a time period of approximately seven months, computer programs were written that contained close to 33,000 instructions and associated documentation of more than 4,000 pages, a fairly ambitious undertaking. It is a fair assumption that it has not been duplicated in many installations. Therefore, although admittedly the programs needed much improvement, the above facts must be taken into consideration when evaluating the work of the contractor.

Since MARC was a pilot, one aim of the project staff was to develop an efficient and accurate system for preparing the bibliographic information for input by experimenting with different procedures and techniques. From the start it was assumed that some form of editing would be necessary, and a series of procedures was designed to prepare the data and to check both the preparation and the conversion of the data for accuracy. At the same time, staff members from several divisions of the Library of Congress were selected to edit the data, and the project staff attempted to determine what kinds of previous experience produced the most successful editors.

As a result both of changes made in the editing and verifying procedures during the course of the experiment and of the variety of backgrounds and proficiencies among the editing staff, large numbers of incorrectly edited and inadequately verified records entered the MARC master data base. By measuring the error rate on verified records, however, it was eventually possible to lower the error rate and to set standards for accuracy of the edited records. No attempt was made to reverify every record input during the first year of the experiment, since over 17,000 records had been distributed by July 1967. After that time both the staff requirements and the editing and verifying procedures were stabilized and the records that were produced and distributed were of somewhat higher quality. Nevertheless, more work remains to be done. The input procedures for the MARC Distribution Service have been designed to further reduce the number of errors.

The account above has been written not as an excuse but to remind the reader that the MARC project was a pilot—a laboratory in which to learn.

The participants' experiences, as described in the following reports, suggest certain formal conclusions. Support for the project by library management is essential. Time must be provided for developing administrative and technical plans, and funds must be allocated early enough to insure the timely acquisition of sufficient staff and equipment.

Close cooperation between computer and library personnel is a prime requirement for implementation of a successful project. Those participants whose computer analysts and programmers were assigned to the library or whose librarians learned to program were most successful. Processing bibliographical information is of such complexity as to challenge highly expert computer personnel. It therefore follows that those assigned must be interested in library problems and strongly motivated toward their solutions through automation.

The Library of Congress Information Systems Office MARC Pilot Project

SUGGESTED TOPICS FOR PARTICIPANT EXPERIENCE REPORT

A. Description of user library and computer configuration used.
B. Original plans for use of MARC.
C. Actual implementation accomplished.
   1. Constraints.
   2. Operating problems experienced.
   3. Results achieved.
D. Reaction to MARC I Format.
   1. Uses made of fixed fields, including language, place, and publisher codes.
   2. Special uses made of variable fields.
   3. Local-use data elements employed.
   4. Modifications made locally to format.
   5. Changes desired in format, with justifications.
E. Reaction to LC-supplied MARC participant programs.
F. Participant experience in relation to computer facility, equipment, programming staff problems, etc. Local programs written for MARC.
G. Administrative or managerial experience with MARC.
   1. Staffing problems; reaction of library staff, etc.
   2. Impact of MARC on local automation plans, budgeting, etc.
   3. Time frame required to implement a local MARC project.
   4. Other organizational or managerial problems.
H. Special studies done by participant.
   1. Performance or use studies made of MARC.
   2. Cost evaluation studies.
   3. Other studies, e.g., statistics of processing times.
I. Distribution problems.
   1. Reaction to air-mailed magnetic tape mode of MARC distribution.
   2. Recommendations on utility of the MARC data base. How it should be organized and maintained by the user; preferences and justification for cumulative vs. non-cumulative distribution; value of the Author/Title file and cross reference tracing files, etc.
J. Conclusions (based on experience with MARC).
   1. Resultant savings, if any (cost implications).
   2. Resultant improvements in library technical processes performance (e.g., faster cataloging), simplification of procedures, etc.
   3. Resultant new products or services not previously feasible.
   4. Other benefits.
   5. Disadvantages.
   6. Summary recommendations to LC.

ARGONNE NATIONAL LABORATORY

Hillis L. Griffin

Information Systems Librarian
Library Services Department

This report concerns our participation in the first year of the MARC project. We feel privileged to have been selected to participate in the project. Our participation has been of great benefit to us as we have developed ways of dealing with MARC data.

Our initial introduction to the operational aspects of the program consisted of two months spent in debugging and modifying the contractor supplied software for the IBM 1401. The software did not work when it was received, and indeed, could not have worked. The source deck of at least one program did not agree with the object deck, which had been modified with patches. Contracting personnel took an extremely pedestrian approach to the problem by ignoring the fact that any 8K 1401 with the ability to read and write magnetic tape must also have the advanced programming package. They used parts of the package but never used the parts which would have made the programs run more rapidly and more efficiently. I would give them extremely low marks for their efforts. We have never had time to test the 360 DOS software.

We based our use of the MARC data more on evaluation of the data than upon trying to integrate the trial project into our actual library operations. If we needed cataloging, the MARC data was available, but we found that in general the proof sheets were available at least two weeks before the MARC data arrived. This fact would indicate that the MARC data must be timely and complete if it is going to be of maximum utility. This timeliness will be especially important as more libraries enter into approval plans with their book dealers and, thus, receive the books almost on the date of publication. We found only two or three occasions for recovery of catalog information from the MARC tapes. I would have liked to have been able to bypass the proof sheets entirely and used MARC as the cataloging base. I think that the preparation of MARC data must be done very promptly as a part of the cataloging operation.

One of the most attractive features of MARC to us was the file containing descriptive cross-references. Libraries have never received this information from the Library of Congress, and each library has been forced to generate this material itself to the best of its ability. The problem in MARC was that the descriptive cross-references were not as important to us, to which they applied. They were treated as an independent file with no ready means of identifying the entries, which could have been of real utility to the cataloging operation. Subject cross-references were not as important to us, since new subjects are noted when the catalog copy is checked against the subject authority file. Descriptive cross-references are sometimes implied, but they take a great deal of time to generate locally. I would hope that the MARC II format will provide this important linkage ability for descriptive cross-references.

In April we assumed the responsibility of distribution of MARC tapes to the users, and I feel that I should recount some of our experiences in this phase of the MARC task since it may be useful in the design of the ultimate MARC distribution system. Tapes were sent to us each Thursday by air freight. The tapes were copied and indexes prepared and the tapes and indexes were sent on to the users on Friday morning. It appears that most users received their materials regularly on Monday morning, and this part of the distribution appeared to be quite successful. We prepared two indexes—one index to card numbers, and another index (in main entry sequence) showed main entry, title, and card number. This latter index operated against the existing File 2, which contained only that information. It appears that some users would have found call number information helpful as a part of the File 2 index, and an adjusted sort (ignoring leading articles, etc.) would probably have been helpful in placing this file in proper sequence. We have several secondary users who receive only this index each week and find it quite useful. This fact might well be considered in the distribution of MARC II, i.e., making the printed index available as a separate service. We attempted to cumulate the indexes monthly, and always at the end of a reel, and did so with moderate success. The major difficulty which we encountered in receipt of the tapes from the Library of Congress was the variable quality of services among the airline freight systems. After several delays with one airline, we switched to another line which has given very good service and eliminated this problem.

A major problem was the lack of experience among the participants in the use of magnetic tape. One user complained about an unreadable tape, and it later turned out that he was trying to read a 7-track tape on a 9-track tape drive. Several users complained of unreadable tapes, but they never reported whether they had taken any measures to rule out machine problems (dirty tape heads, program difficulties, 7-track tapes on 9-track drives, etc.). We never learned where the tape had failed and whether the failure was caused by some physical defect in the tape itself. Some users were unable to cope with problems which are rather commonplace to people who work with computers. They could only take the programs and send them along with the tape to the computer center and hope that all would go well.

I think that librarians or other personnel familiar with basic data processing operations are needed at any installation which intends to use MARC data. The users really don't seem to know what to do when minor problems arise, and they seem more often than not to simply take to the typewriter to let someone (perhaps not even us, the distributor) know that they got a bad tape some weeks back. Of course, the tape has been returned by that time and has probably gone out again in the cycle, and the opportunity to locate and correct this difficulty is long past. One tape was even returned, naked in the box, without its protective plastic reel can. Someone else kept returning the tape enclosed in a slip-ring instead of a reel can. I kept wondering what was so difficult about putting the reel back into the reel can for return shipment. It might also be noted that we cleaned and recertified all of the MARC tapes, in September and October, and several obviously bad tapes were removed from the system.

We had hoped to make the MARC services generally available to several secondary users. One problem which prevented effective use was an inability to generate cards with the call numbers in the upper lefthand corner. We have since written a subroutine which will separate the class number from the Cutter number but this does add overhead to the program and is not entirely consistent, especially with some juvenile books. Perhaps the Cutter number should be delimited or set into a field of its own for maximum utility. Other potential secondary users were not prepared to cope with problems involved in operating the programs, or they had computer systems which were too small to utilize the tapes. We had hoped to generate a variety of output products from the MARC tapes (listing by series, subject, class, juveniles, etc.), but the press of local problems involved in new computer acquisition and installation did not leave sufficient time. I think that these products would have been especially valuable to non-MARC participants. We have also generated a large amount of software to handle the MARC I format on the 1401 and have very nearly completed their reduction to callable subroutines to be added to an Autocoder library. One of the problems which we had, for example, was in interfacing the MARC character set to our IBM "H" print chain. This required that we translate the parenthesis, brackets, colon, and semicolon, question mark, apostrophe, quotes, and exclamation point to usable substitute characters (the parenthesis themselves were coded for the 360 rather than the 1401). We also had to delete diacriticals. Suprisingly, this problem, properly approached, did not really slow down our printing cycle as much as might be imagined. In a distribution system, however, some consideration might be given to providing chain-adjusted tapes if a sufficient number of users request them.

I have taken every opportunity to explain the MARC program to other librarians, and we have had a number of visitors and letters requesting more complete information about the MARC program. While this has taken time, I feel that it has been time well spent, for I think that we need to make *all* librarians aware of the great potential of MARC for their libraries. We have been pleased with the MARC tapes and given timely and complete tapes (which we realize was impossible in the pilot project), we would have no difficulty in using them in our technical processing activities. We very definitely want to see MARC grow and prosper, and we hope that it will become an important aspect of the program of the Library of Congress.

GEORGIA INSTITUTE OF TECHNOLOGY

John P. Kennedy
Data Processing Librarian

A. User Library and Computer Configuration

The Georgia Institute of Technology's Price Gilbert Memorial Library is an important information center for science and technology. Outstanding collections in the fields of engineering and science have been developed to support graduate study and research. On July 1, 1967, there were 764,542 library items, including 422,367 volumes, 239,843 microtexts, and 102,332 miscellaneous items, such as maps, slides, films, phonorecords, and pamphlets. Approximately 10,000 serials, including 4,000 periodicals, are received currently. About 75 percent are in scientific and technical fields. The services of the library are rendered by a staff of over 60 persons. Institutional support is excellent with the budget for the current year approaching one million dollars.

The Rich Electronic Computer Center's Burroughs B–5500 computer has been used for the MARC project. The configuration of the B–5500 includes:

- 2 Central Processors
- 8 Core Modules, each with 4,096 forty-eight bit words
- 8 Data communications telephone line adapters
- 3 B–475 Disk Modules with a total of 28.8 million characters of storage
- 4 I/O Channels
- 10 B–442 Magnetic Tape Transports, 200 and 556 bits per inch, 7 level
- 8 B–124 Card Readers, 800 cards per minute
- 2 B–321 Line Printers, 600 lines per minute
- 1 B–304 Card Punch, 300 cards per minute

B. Original Proposal for Use of MARC Data

We are currently converting to the Library of Congress classification system. Our plans call for the printing of book catalogs beginning in the summer of 1966 to include all materials classified by the LC system. Initially we will also add cards to our centralized card catalog for these items. The catalog cards will be prepared on an automatic tape typewriter, and a byproduct tape will be converted to punched cards for computer input. Maintenance of both the card catalog and book catalog for a time will permit comparative studies of the relative usefulness and costs of the two. (A decision will be made later as to which of the two forms should be maintained permanently.)

If we are able to participate in the MARC project, we will design our record formats to be as nearly identical as practical to those used in the MARC project. Machine-readable catalog copy would be used for computer input in place of the byproduct data from the card production operation. A printed listing of the machine-readable records would be utilized by the cataloger to determine whether machine-readable copy is available. It would then be necessary for us to keypunch only the identification number of the machine-readable record, shelflist data such as special locations and number of copies, and any changes in the LC copy which were absolutely necessary. Our present policy is to accept the complete LC call number and cataloging data, so the number of changes required should be minimal.

We have already begun to modify some LC entries slightly in order to facilitate machine filing. If the MARC project uses LC entries without working out a filing system, we would modify some entries for filing. The required entries for the various catalogs would then be developed from the master record, and catalog cards would be printed. The new entries would then be merged with our book catalog files for the printing of catalogs and supplements.

C. Implementation Accomplished

Our original plans for use of the MARC data are being carried out, although at a somewhat slower pace than anticipated. The only change from the plans outlined in the original proposal is that we are not altering entries in any way for filing.

Implementation of the MARC system has caused little disruption of existing procedures. When a book is received in the Cataloging Department, a clerk checks the printed MARC listings to determine whether or not a record for the title is available on the MARC file. Checking is done by LC card number which is usually available either from the book or from one of the sources used for verifying the order data. If the cataloger has cataloging copy available in the form of an LC deposit card or a copy of the published NUC entry, she makes any needed additions or changes on the copy and sends it on to the Data Processing Department.

The LC card number, several local use data items, and any changes or additions to be made to LC cataloging are punched. The punched cards are used to select the required records from the MARC file. The same program converts the character codes to those used by the B–5500, eliminates shift codes, converts the records to our local format, makes any indicated changes or corrections, and prints the records for proofreading. The printed output is proofread against the LC deposit card or copy of the NUC entry, if either has been sent along to Data Processing. Otherwise the proof listing is returned to the cataloger for checking against the book. Changes are punched and run against the tape file of selected records. The corrected tape output is used to print catalog cards and input into the book catalog subsystem. The correction program also produces punched card output which can be converted to paper tape and used to produce book cards, book pockets and spine labels on a Flexowriter. The punched card output has not been used except for test purposes at this time.

The card print program provides the option of printing the cards in sets for each title or in presorted and alphabetized order for each of the catalogs into which they will be filed. Thus far the cards have been printed in sets so that they may be conveniently checked by the catalogers. As our confidence in the card print program is increased, we expect to eliminate this final check by the cataloger and to print the cards in the order required for filing.

The format of the cards produced by this system varies from the conventional card format. The main difference is that the collation is printed as the last element of the title paragraph rather than as a separate paragraph. This practice was adopted as a space saving device both on the cards and in the book catalogs. A few other variations from conventional card format have been adopted for the same reason. The cards produced are not unit cards but vary according to the requirements of the various catalogs. Cards for the Union Catalogue of the Atlanta-Athens Area, for example, omit the call number, tracings, and all notes except series notes, but include a symbol for the Tech Library. The makeup of sets for titles going to different locations also varies. Each title cataloged for the Architecture Library produces a complete set of cards for the main library, as well as a set for the Architecture Library, while only an extra shelflist card is produced for titles going into the archives collection or to the Southern Technical Institute.

The tape used as input to the card print program is also used to update our master file and generate entries for the book catalogs. A tape produced from the Flexowriter card production subsystem provides a second source of records for the master file and book catalogs. Finally, the print entries generated are sorted, and the book catalog supplements and new book list are printed.

We began using the card print program the last week of May 1967, and by the end of November had printed 13,715 cards for 1,752 titles. Normally, titles are batched and punched once a week. The elapsed time from punching of the selected cards until the completed catalog cards are returned to the catalogers has normally taken about one week. The programs to produce the book catalogs, however, have not yet been used on a regular basis. We have produced several test printouts and a pilot edition including entries for about 2,400 titles, but a number of program revisions remain to be completed.

D. Reaction to the MARC I Format

1. Local uses of fixed fields.
   a. The LC card number and supplement number are used for maintaining the file order and for selecting records from the MARC file.
   b. The Type of Main Entry is used in setting up sort keys for the book catalog entries.

c. The Language 1 field is used in preparing the sort key for title entries.

d. The Type of Publication Date, Date 1, and Date 2 are used in setting up the heading on extension cards and in setting up sort keys.

e. The Length of Record is used in the character and format conversion of the record.

2. Special uses of variable fields.

No special use has been made of the variable fields.

3. Local-use data elements employed.

a. Location of the item in our library.
b. Cataloger's and keypuncher's initials.
c. Date selected for our file.
d. Code to show method of acquisition.
e. Indicator for titles reclassified from the Dewey Decimal classification.
f. The call number and the card number are carried in a field edited for sorting as well as in a field edited for printing.

4. Local Modification to the MARC Format.

We have modified the MARC format in order to facilitate programming and to decrease running time for our programs which are all written in COBOL. In the COBOL language a series of variable length fields can be individually manipulated only by subscripting through, character by character. In order to avoid the need for handling the data character by character each time it is printed, we have formatted the variable length data elements in paragraph form, carried in a variable number of fixed length segments. In our print programs each segment becomes a print line. The segmentation is done at the same time that we subscript through the data for code conversion. The data need not be handled character by character again unless we wish to alter the length of our print lines. The use of a variable number of fixed length segments produces a record which is slightly longer than when the corresponding data is carried in strictly variable length fields. We believe the decreased running time more than compensates for the increased record lengths.

A second change which we made in order to facilitate programming has been the use of a consolidated directory. Instead of beginning each variable length field with a tag and the field length, we use a consolidated directory giving the tag and length of each variable length data element.

5. Changes desired in format.

We approve of the modifications made in the MARC II Format, especially the use of a directory. We recommend that in the new format the use of shift codes or escape codes be avoided in fixed lengths or semifixed length fields. The shift codes which occur in almost all of the fixed fields in the MARC I format serve little purpose since the fields are seldom used for printing. Since the occurrence of shift codes is not always predictable, their use makes the fixed length fields variable in length. The occurrence of shift codes in the LC card number has meant that we must eliminate them from all records received before selecting any records from the file.

We hope that character counts for record lengths and field lengths on 7-channel tapes will be true character counts, including all shift codes or escape codes. The starting character positions in the directory for 7-channel tapes should also indicate the actual starting position on the 7-channel tape rather than what would be the starting position if 9-channel tape were being used.

We also suggest that a code for the language of the title would be helpful. Occasionally the language of the title differs from the language of the work, and these cases can cause difficulties in filing if there is no indication of the language of the title. An indicator might be set in those cases in which the language of the title differed from the language of the work and in those cases only, a code for the language of the title supplied.

E. Reaction to LC-Supplied Programs

No programs which were supplied by LC could be used on the B–5500.

F. Experience with Computer Center

Our Rich Electronic Computer Center has given excellent support and cooperation throughout the MARC Pilot Project. The Library has not been billed for any computer time used in the project or for a substantial part of the programming time spent on the project. We have also received priority handling for many of our MARC runs on the computer.

The use of the Burroughs B–5500 to process tapes written on another manufacturer's computer presented no serious difficulty. Since the machine codes for many of the special characters differ, we found it necessary to scan the data character by character and to look up each character in a conversion table. For the apostrophe which is not available on our printer we had to substitute another character. All shift codes and all codes for diacriticals were eliminated in the conversion routine.

Our most serious problem in using the B–5500 has been the quality of the printed output. The B–321 Line Printers used on our B–5500 are drum printers which tend to produce wavy print lines when not in exact adjustment. Two batches of cards have had to be reprinted because of the extreme waviness of the printing. Recently, we have tried to have the cards printed immediately after the Burroughs engineers have adjusted the printer in their regular daily maintenance. The quality of the printing has been improved but still leaves much to be desired.

We have also experienced some difficulty in getting all of the operators to align the forms with the required degree of precision. In order to get maximum data on the cards we have allowed minimal margins and the forms must be aligned carefully. Our operators are not accustomed to the need for high quality output, but this problem is being overcome.

We have usually had overnight turnaround time on our programs. This has been satisfactory for production runs but has resulted in considerable delay in testing and correcting programs.

A description of the programs which we have written for use in the MARC Project follows.

1. The MARC MAINTENANCE program updates the MARC file from the weekly tapes supplied by LC.
2. The SELECTION program selects records from the MARC tape, converts character codes, and sets up a record in a format adapted to our use. Data items may be added, changed, or deleted from the MARC records selected. Two prooflists are printed: one for records for which cataloging copy has been checked by the cataloger and a second for records for which cataloging copy has not been checked.
3. The CORRECTION program provides for further editing of the records written by the SELECTION program. A proof listing is printed for those records being changed. This program punches cards which can be converted to paper tape and used to type book cards, book pockets, and spine labels on a Flexowriter.
4. The CARD PRINT program prints catalog cards for the selected records. The cards may be presorted and alphabetized for filing.
5. The FLEXOWRITER program processes output from our Flexowriter subsystem for records which are not on the MARC tape, as well as keypunched changes, corrections, and deletions for the master file and book catalogs.
6. The UPDATE-GENERATE program merges and sorts to call number order the MARC and non-MARC input for the master file. It updates the master file and generates entries and changes for the print files. It sorts the entries and changes by catalog and by the generated sort fields.
7. The WEEKLY PRINT program prints the weekly supplements for the three printed catalogs and a list of new books in call number order. The weekly supplements cumulate until the next monthly supplement is produced.
8. The MONTHLY PRINT program updates the print files for the catalogs and prints monthly supplements which are cumulated until the next printing of the full catalog. A new book list for the month is also printed.

G. Administration Experience

1. Staffing problems.

We have encountered difficulty in obtaining programmers with experience in programming of the type required for library work. Although two experienced COBOL programmers were assigned to the project, one of them produced no usable coding. After a period of 11 months in which almost six months of his time was charged to the project, we reached a mutual agreement that his work was not productive and no more of his time should be devoted to this project. Even though he was a capable and experienced programmer with several years experience programming accounting systems, he had little inclination or aptitude for programming a system with the complexities, uncertainties and variability of data inherent in this project. Hence, almost all programming for the project has been done by a single programmer who does have unusual aptitude and ability for this type of work. She was working on other assignments at the time the MARC project began, however, and has been unable to devote all of her time to the project.

The reaction of the Library staff to the MARC project has been one of pride and enthusiasm. Our catalogers have been interested in finding ways to facilitate the operation of the system and have adapted their procedures to meet the requirements of the system without complaint or undue difficulty.

2. Impact on Local Automation Plans.

Our plans for automation now center around the use of MARC records and the MARC format. Specifically, we expect to modify our local format to conform to the MARC II format as far as it is practical while still programming in COBOL. The only variation from the MARC II format that we think we may wish to retain is the use of fixed length segments rather than variable fields for the variable length data items. We are also considering the possibility of having our catalogers assign MARC codes and indicators and fixed field data for locally cataloged items.

We expect to begin work soon on system design work to include acquisitions procedures in our MARC system. We will delay further work on a serials system until the content of serials records in the MARC II format is announced by the National Serials Data Program. We expect to use the MARC II format for files in systems which we design hereafter.

3. Time Frame for Implementation.

The time required for implementing the system greatly exceeded our expectations. We were notified of our selection for participation in the project in February 1966 and hoped to be ready to begin operation in September 1966. Actually we did not produce our first usable catalog cards until the end of May 1967. The first pilot edition of the book catalog was produced in November 1967, but several improvements to the programs which produce the book catalogs are still underway.

The time required for the implementation of a local MARC system depends on a number of factors, including the complexity of the system undertaken, the availability of systems and programming talent, and convenience of access to the computer. In our case, the system which we are implementing includes both card and book form output and entails the maintenance of a master file of the Library's holdings of monographic titles. Another complication has been the requirement that the system handle input from our Flexowriter card production subsystem as well as from the MARC tapes. Programming for the processing of Flexowriter input has been included as a part of our local MARC project, and this effort is included in the time given for programming the system. A final complication in our system was the necessity for programming conversion of codes on the MARC tapes to those used by the B-5500.

A lack of knowledge of the details of the MARC format resulted in some delay during the first eight months of the project. The initial report on the MARC format was received in April 1966 and we proceeded with programming on that basis. The final report giving all of the details of the format and a test tape were not available until October 1966.

Implementation of our system has also been delayed by the difficulty in obtaining qualified programming help. We estimate that 32 months of work was devoted to system design and programming through the end of November 1967. Our primary programmer has worked 14 months full time on the project during a period of 21 months. The project director spent approximately 12 months on system design and programming during the same period. As stated above, one of our programmers spent six months in unproductive effort before leaving the project. We believe that two qualified programmers could complete programming for a comparable system in approximately one year of full time effort.

H. Special Studies by Participant

1. Performance or use studies.

During the months June through November, we cataloged 4,817 English language monographs. Records for 40 percent (1,929) of these were available on the MARC tape at the time of cataloging. This is an average of 74 titles per week. The number of titles has been gradually increasing. The average number of titles selected from the tape each week from June through August was 62, while the average number each week from September through November was 87.

We found that the MARC records usually arrived promptly enough to meet our needs. Our policy has been to hold books for cataloging until LC copy, if available, is obtained, unless the book is designated "Rush." We have continued to hold books for cataloging until either the LC Title II deposit card or the MARC record arrives. We made a check to see if records were later added to the MARC file for 75 titles which we processed through our Flexowriter system because MARC records were not available at the time of cataloging. These 75 English language monographs with 1967 imprints were cataloged during the first week of October with the use of cataloging copy from LC deposit cards or from the published NUC. One of these titles appeared on the MARC tape dated October 5 and another was on the October 12 tape. None of the other 73 had been added when the November 30 listing of MARC records was checked. We have concluded that records which are selected for MARC processing are added to the file at about the same time that we received the LC deposit cards and that this is usually prompt enough for our needs.

2. Cost evaluation studies.

The cost of computer time for maintaining the MARC file, selecting, converting, reformatting the records for our use, making required changes and corrections, and printing catalog cards has averaged about $.30 per title during the months of October and November. Computer time is figured at a rate of $140 per hour for process time and $47 per hour for I/O time. We have averaged eight cards per title for MARC records, including cards for deposit in the NUC and the Union Catalogue of Atlanta-Athens Area, duplicate sets of cards for titles cataloged for our Architecture Library, and extension cards. The average machine cost per card is about $ .035.

Costs for catalog card stock, tab cards, and a prorated part of card cutter and keypunch costs average about $.015 per card or $.12 per title. We estimate that we have been spending about 14 hours of professional time and 2 hours of clerical time for each 100 MARC titles processed. At $4.00 per hour for professional time and $2.00 per hour for clerical time this gives a labor cost of $.60 per title. The total cost per title is then about $1.02 or just under $.13 per card.

We expect the labor costs to drop dramatically as the processing of MARC records becomes more routine and program defects are corrected. Another reason for the great amount of time spent proofreading, however, has been the number of errors in records on the MARC tapes. In 988 records selected in the last eight runs, we found and corrected 120 errors, or almost one error for every eight records. A decrease in the number of errors on the records added to the file will naturally reduce the amount of time required to proofread and make corrections.

We are now training a nonprofessional technical assistant to proofread and make corrections for MARC records. We hope to decrease the time spent in these activities to not more than five hours per 100 titles. This will decrease costs to about $.60 per title and $.075 per card. To this figure, however, must be added the price of tapes when they are available on a subscription basis. Our present card production cost is lower than costs for using LC printed cards or for cards produced locally on the Flexowriter with a byproduct machine-readable record.

The cost of computer time used in the development and operation of the system through November 1967 was $4,125. Personnel expenses for system development amounted to $24,496. The total cost to the Library and the Computer Center for the developmental work was $28,621 through November 1967.

I. Distribution Problems

1. Reaction to the mode of distribution.

The use of air mailed magnetic tapes has been satisfactory to us. We usually receive a tape dated Thursday the following Monday or Tuesday.

2. Organization of the file.

We maintain our MARC file in card number order and recommend that this organization be continued. The cost of our file maintenance and selection runs has been gradually increasing as the MARC file has grown. We anticipate that it will become more economical to divide it on a chronological basis. Probably only those records received in the last 12 to 24 months will be retained in the active file that is processed weekly. The file of older records will be processed only as often as the number of selections to be made from it warrants.

Cumulation by LC until a reel is filled has been very helpful. This has made it possible to minimize the number of file maintenance runs. Since most of the records to be selected each week are already on our cumulative MARC file, we do not update the file each week but wait for several weeks until we begin to get an accumulation of records to be selected from recent weekly additions. In short, fortnightly, or even monthly, distribution of the tapes would probably suffice as long as we continue to receive on a weekly basis a listing of the new records added.

The cumulative listing of new records by card number is essential to our system; the author/title index is helpful. If the listings were discontinued, it would be necessary to update our MARC file ourselves and produce a cumulative card number listing each week. In this case, the cumulation by LC of records on a reel would be of no value to us.

We have made no use of Files 2, 3, and 4. File 2 would be useful if the author/title listing supplied by LC were discontinued. We anticipate no use of File 3 since we utilize the printed list of LC subject headings as a supplement to the subject section of our card catalog rather than include typed cross-reference cards in the catalog itself. File 4 might be of some value, but we have no plans for it.

J. Conclusions

We regard the MARC Pilot Project as highly successful. Even though the implementation of our system required much longer than expected, we have now realized or are close to realizing all of our planned goals. We have been producing catalog cards at a cost lower than the cost of other methods of production in this Library. About 14,000 cards produced from the MARC records have been filed in our catalogs or sent to union catalogs without our having received any objections to the use of a limited all-uppercase character set or to revisions in format.

We have produced a pilot version of a book catalog containing entries for approximately 2,400 titles selected from the MARC tapes or processed through our Flexowriter system. While this catalog is not yet ready for public use, we see no serious problems which cannot be corrected as we refine the programs. Our catalogers have been enthusiastic in their acceptance of the MARC system. The use of MARC records, along with the introduction of several other simplifications in cataloging procedures, has permitted us to process approximately twice as many titles during the period July through November 1967 as during the same months of 1966. This has been accomplished with no increase in the staff of the Cataloging Department, although the Data Processing Department has contributed about 16 hours per week to card production work.

We recommend that the MARC project system be converted to a permanent program and expanded in its coverage as rapidly as possible. We regard expansion to include all English language monographs and serials as especially desirable.

We regard the proposed MARC II format as an important improvement over the MARC I format. We expect to reprogram for the MARC II format and to revise our own tape format to be identical, or nearly identical, with it. We suggest that consideration be given to making available a version of the MARC tape with no shift characters or escape characters for those users having printers with a small character set. Conversion of the tape to eliminate diacritical marks and shift characters at LC could save programming effort and processing time for these users. We recommend that fixed fields and especially the control number should not contain any shift characters. Character counts for 7-channel tapes should be true character counts for the 7-channel tapes rather than character counts for 9-channel tapes.

We also hope that the Library of Congress will continue to supply the cumulative listings of card numbers along with the tapes. Again, this would save programming effort and a considerable amount of processing time each week for many users.

Finally, we recommended that great care be given to the editing of MARC records. We have been disappointed to find that the frequency of errors in the MARC records necessitates a considerable amount of proofreading. We understand that in the pilot phase of a new system the number of errors will inevitably be high, but we hope that when the MARC program is put on a permanent basis the accuracy can be improved to at least equal that of the printed cards.

We realize that several of these recommendations would entail extra effort and costs at the Library of Congress and might result in slightly higher subscription costs for the tapes. The extra time and effort at LC, however, could save a comparable amount of time for each of many users, and the extra cost would then be spread among these users.

HARVARD UNIVERSITY LIBRARY

Foster M. Palmer
Associate University Librarian

A. Environment

The Harvard University Library has 7.8 million volumes, housed in nearly 100 different libraries, of which the Widener Library is the largest with 2.4 million. The Library as a whole serves a teaching and research staff of some 7,000, over 15,000 students, and nearly a thousand other readers, principally visiting scholars and faculty members of nearby institutions.

Almost simultaneously with the beginning of the MARC project, the Harvard University Computing Center opened a substation in the Widener Library. The equipment available in this substation, which was established in space provided by the Library in return for the convenience of having a computer close at hand, consists of an IBM 1401 with 8,000 positions of core storage, four 7330 tape drives, a 1402 card read-punch, and a 1403 printer with 120 character print chain capability, giving the possibility of printing in uppercase and lowercase. The computer has all the usual special features such as high-low-equal compare, sense switches, and advanced programming, as well as some not so common, including space suppression, column binary, and multiply and divide. Although this computer is located in the Library, it is operated by the Computing Center, and the Library pays the same charges as other users, $30 per hour at present and during most of the period covered. The Computing Center's main installation, less than half a mile distant, has two IBM 7094's to which tapes can be sent for sorting.

B. Original Plans

The principal specific study envisioned at the beginning of the Library's participation in the MARC Project was an evaluation of the promptness with which data were received. To this end, extensive records were kept throughout the greater part of the year 1966-67 of the ordering, receipt, catalog searching, and filing dates for books thought to be within the scope of MARC. Somewhat later the Library received a grant from the National Science Foundation, the larger part of which was to be used to undertake the creation of records in the MARC format. Beyond these specific projects, it was hoped to gain experience in handling machine-readable bibliographical records supplied by a central source and to seek ways of making this information conveniently available to catalogers.

C. Constraints, Problems, Results

Constraints may be considered at various levels. The major one was that the Harvard University Library is a very large and very old ongoing operation whose basic processes are not tampered with lightly; introduction of a mechanized system on which actual production depends cannot precede extensive testing. At a more specific level, it was a disappointment that, although the printer was equipped for uppercase and lowercase operation, the print chain actually available was not the FAU chain developed for library use, but a Courier or Text 90 chain which does not include diacritical marks.

Given the limited goal of experimentation rather than production, operation problems were not severe. It was recognized in part before the first tapes were received and more fully immediately thereafter that some of the standard programs furnished by the manufacturer and by the Computing Center were unsuitable for use with MARC, and special programs, even of a utility nature, would be needed. Specifically, a tape dump that would give proper character numbering for tapes written with word marks was required. Furthermore, a special merge had to be written for MARC, since the merges supplied by the manufacturer for an 8K machine would not come close to handling records 2,008 characters long. More importantly, a series of macros was written. These are more fully described in the section on programming.

When the tapes actually started coming, there were various small problems at the beginning, but these were surmounted and a cumulated local file of MARC records was built up, with date of receipt added in the local use portion of the fixed field. Availability of the MARC data was widely announced within the University Library, but requests were few at the start. This was not surprising because the number of titles represented at first was small, and all were in English, not an area with a particularly high incidence of problems. This is not said to denigrate the choice of English for a start, a natural and almost necessary one, but to suggest why initial demand was so low. The first actual uses of the tapes were when bibliographic listings of all MARC records for books classed by the Library of Congress in R were run off for the Countway Library of Medicine and somewhat later all those classed in B were printed for the Andover-Harvard Theological Library.

The program to select titles by the first letter of the classification was very simple, but it was obviously also desirable to be able to select on variable fields. However, a fully generalized program of this nature seemed impossible because of the infinite variety of possible selection criteria. The solution was to write what might be called a skeleton selection program into which *ad hoc* extensions were inserted, as described more fully in the section on programming. As an exercise, an extension was written to select books published in or relating to India. The books published in India were found by a simple examination of the fixed field, while those relating to the country were found by a combination of techniques, including examination of the classification number and occurrence of the word India or Indian in a corporate heading, the title, or a subject heading (except that works containing the word Indian but classified in American history were assumed to relate to the American Indian and rejected). Although this program would fail to select a work whose classification fell in a special subject rather than in the general class for India, and whose subject headings and title contained such terms as Bombay or Uttar Pradesh but not the word India or Indian, it still gave fairly impressive results, and the Library's South Asia book selection specialist has requested an elaboration of it to include certain neighboring countries and additional segments of the classification. Perhaps the most interesting spontaneous request for a specific selection was from the Music Library, which subscribes to the Library of Congress proof sheets for the music class (M), but wondered what other music might be represented on the tapes. A search was run looking for the word "music" in the collation but excluding words classed in M. The results were quite interesting and useful, including a number of works on folklore and liturgy and one on the history of printing which included a section on the history of printing of music. Also retrieved was Hawken's *Copying Methods Manual*, since it includes music among the material illustrated. More recently, the Harvard Law Library has become interested in the manner in which the Library of Congress is applying the newly activated K class, and has requested that listings of works classified in K be supplied on a regular basis.

Very early it was decided that the author-title listing supplied with the MARC tapes was unsatisfactory, and steps were taken to provide one that would be more useful. This and related problems turned out in fact to be the major focus of Harvard's participation in the MARC Project. The author-title list as supplied suffers from two main deficiencies, one of them common to many library lists, the other peculiar to machine processed material. The first is the simple fact of limitation to a single listing under main entry. While this is usually an obvious and highly useful heading, in a very substantial minority of cases it is an assigned entry which may properly be considered "main" in the context of a catalogue, but is not necessarily the most important entry for finding purposes. The second and more glaring deficiency is that a raw computer sort is used to arrange these main entries, with the result that many works with title entry appear under an initial article; the first entry on the first alphabetical list was for an anonymous work called *A little pretty pocket-book*, and the first on the current list is *A cabinet of curiosities*. Furthermore, the names of authors whose initial letters have diacritical marks come to the very head of the list, a phenomenon first noted the third week, when František Šorm led off, filed as 8SORM. Diacritics on other early letters, while less dramatic in their effects, led to filing poor enough to cause items to be missed.

The Harvard solution to the first problem was always to make a title entry on the ground that title is the most stable element in the description of a book, particularly a book one has in hand. Next to title, and perhaps even preferable to it when simple declared authorship is involved, is a personal name associated with the work. The Harvard index included personal names as main entries and those traced under tag 71 as added entries. It would also have included personal names as subject entries had these been distinctively tagged. The initial plan was not to include corporate entries at all, but since corporate authors are often not repeated in the title statement even of very short and non-distinctive titles, this proved to be a mistake and the program was patched to print them as well. The index, which printed two-line entries, has been run off from time to time and placed in the Catalogue Department.

Making multiple entries was intellectually simple although of course it increased printing time. Improving the filing was a much more complex question. It was not difficult to strip off articles, particularly so long as the corpus was virtually entirely in English (although an interesting problem was pointed out when a book in English, and coded as in English, had a title beginning with a Spanish article—*La vida*, by Oscar Lewis). The proper machine filing of personal names proved to be a much more difficult matter and one that is not yet fully resolved, although much progress has been made with surnames.

To illustrate this work, inspired by the MARC Project and made possible by having a moderately large data base containing a wide variety of names illustrating actual problems, a *Personal Name Index to the MARC Tapes* was produced (see Figure 1). Fortunately MARC

```
Harvard University Library Name Index to the MARC Tapes, January 24, 1968                                                  Page 134

O'Grady, Desmond                  66-18017 | Oliver, Michel                   AC67-2335 | Oncken, Hermann                   66-24743
O'Hara, Frank                      67-4141 | Oliver, Peter                     67-28255 | Onclin, Willy                     67-28693
Ohara, Houn                       66-28571 | Oliver, Raymond                   66-22340 | O'Neal, Robert                    66-31378
O'Hara, John                      67-12717 | Oliver, Raymond                     67-368 | O'Neil, Robert M.                  67-6946
O'Hara, Kenneth                   67-94219 | Oliver, Robert W.                  67-6633 | O'Neill, Barbara Powell           66-21160
O'Hara, Patrick                  67-104021 | Oliver, Roland Anthony           67-110448 | O'Neill, Charles Edwards          66-21529
Ohdedar, Aditya Kumar            SA66-7744 | Oliver, Stephen J. L.             67-70793 | O'Neill, Eugene Gladstone         67-15461
O'Higgins, Brian                   67-2084 | Oliver, William                   66-26351 | O'Neill, Gil                       66-9087
Ohlin, Bertil Gotthard            67-17317 | Olivier, Daria                    66-21736 | O'Neill, Hugh                     67-98462
Ohlin, Goran                      67-81274 | Olivier, Daria                    66-78653 | O'Neill, J. C.                    66-78535
Ohlsson, Ib                       67-15962 | Ollard, Richard Lawrence          67-24063 | O'Neill, Mary (Le Duc)            AC67-246
Ohlsson, Ib                     AC66-10727 | Olle, James Gordon Herbert        67-85940 | O'Neill, Mercia                    66-9087
Ohlsson, Ib                      AC67-2511 | Olley, R. E.                      66-27851 | O'Neill, Michael J.               67-10030
Ohmer, Merlin Maurice             66-14621 | Olliver, J. G.                    67-83868 | O'Neill, Richard W.               66-21160
Ohno, Susumu                      67-16669 | Olmsted, Frederick Law            67-14738 | O'Neill, Robert John              67-11678
Oikonomides, Al N.                65-26123 | O'Loughlin, Carleen               66-29598 | O'Neill, Trevor                   67-88224
Oinas, Felix J.                   66-63527 | Olschewski, Alfred                67-17546 | O'Neill, William L.               66-11869
Ojha, Gopesh Kumar               SA67-5517 | Olsen, Arthur Robert              67-14242 | Ong, Walter J.                    67-10575
Oji, Agwu                         67-82395 | Olsen, Eigil                      66-78520 | Ong, Walter J.                    67-24508
Ojo, G. J. Afolabi                67-75209 | Olsen, Einar A.                   67-10700 | Onin ki                           66-14595
Okam, Eric C. N.                  67-82050 | Olsen, Evelyn Guard               66-24419 | Ono, Jo                            67-1152
Okamura, Koson                   AC67-2229 | Olsen, Jack                       67-10959 | Ontario                           67-97734
Oke, K. H.                        66-29596 | Olsen, Oliver Wilford              67-4202 | Opdahl, Richard D.                67-14085
O'Keefe, Marian K.                66-30461 | Olsen, Oluf Reed                  66-77416 | Opdycke, John Baker               66-26000
O'Keefe, Robert                   67-61049 | Olsen, Robert Arnold              67-65359 | Opie, Iona (Archibald)            67-88054
Okereke, Onyemaechi               67-82396 | Olsen, Theodore V.                67-10357 | Opie, Peter                       67-88054
Okey, Loren La Mont               67-23464 | Olsen, Thomas Marshall            66-30657 | Opler, Marvin Kaufmann            66-27935
Olafson, Frederick A.             67-16038 | Olshan, Nathan H.                 67-60383 | Opler, Marvin Kaufmann            66-27935
Olby, Robert Cecil                66-23853 | Olson, Charles                    66-27613 | Oppenheim, Abraham Naftali        67-86839
Olby, Robert Cecil                66-29597 | Olson, Clair Colby                66-17748 | Oppenheim, Joanne                  67-4377
Old time vendor                    67-6844 | Olson, Gene                       67-10613 | Oppenheim, Joanne               AC67-10436
Oldeberg, Hans                    67-90708 | Olson, Gene                       68-10139 | Oppenheimer, Harold L.            66-24397
Oldenberg, Otto                   67-10877 | Olson, Gene                       AC67-986 | Oppenheimer, Jane Marion          67-14098
Oldenburg, Claes                  67-16292 | Olson, Harry Ferdinand            66-28730 | Oppenheimer, Samuel L.            66-24955
Olean, Sally J.                    67-8294 | Olson, Harry Ferdinand             67-2932 | Opper, Frederick Burr             67-13534
O'Leary, C. O.                    67-93170 | Olson, Harvey Stuart              66-16661 | Oppersdorff, Mathias T.            67-6595
O'Leary, Erica                     67-3497 | Olson, Harvey Stuart              67-11309 | O'Quinn, Garland                  67-21321
O'Leary, Frank                    67-17330 | Olson, Herbert Vincent            67-24853 | Oraison, Marc                     67-13759
O'Leary, Frank                  AC67-10086 | Olson, Lawrence Alexander         67-15754 | Oravetz, Jules A.                 67-17396
O'Leary, John Gerard              67-85951 | Olson, Mildred Thompson           66-20840 | Orbaan, Albert                   AC67-1280
O'Leary, John Gerard              67-87530 | Olson, Reuben M.                  66-20528 | Orbaan, Albert                   AC67-3987
O'Leary, John Gerard              67-93526 | Olson, Ronald LeRoy               67-16032 | Orbell, Albert                    67-75027
O'Leary, Michael                  67-18985 | Olson, Ronald LeRoy               67-16032 | Orchard, Richard E.               66-26784
O'Leary, Michael Kent             67-18274 | Olson, Theodore A.                67-20266 | Orchard, Vincent                  66-25570
Oleck, Howard Leoner               67-9241 | Olsson, Naima                     67-95554 | Orchin, Milton                     67-5208
Olela, Henry                      67-81290 | Oluwasanmi, H. A.                 66-69196 | Ord-Hume, Arthur W. J. G.          67-9564
Ol[illegible]en, Henning Lind     65-22183 | Olyslager, Piet                   67-94248 | Orde-Browne, Granville St. John   67-81514
Olgin, Joseph                     67-10460 | O'Malley, Michael                 66-20288 | Ordway, Frederick I.              66-22417
Olin, Harold Bennett              66-23400 | O'Malley, Raymond Morgan          67-94093 | O'Rear, Frankie (Richmond)        67-23485
Oliva, Leo E.                     67-15579 | Omar Khayyam                       67-6842 | Orear, Jay                        66-26754
Oliva, Peter F.                   67-12078 | Omar Khayyam                      67-17906 | O'Rear, John                      67-23485
Oliveira Lima, Manuel de          66-24742 | O'Meara, Thomas A.                67-93522 | O'Relly, Edward                   67-31942
Oliver, Egbert Samuel              67-6487 | Omholt-Jensen, Edvard            67-108385 | Orentlicher, Herman I.             67-1362
Oliver, Eloise McLean             67-22995 | O'More, Peggy                      66-9244 | Orewa, G. Oka                     67-92641
Oliver, Jane                      67-94830 | O'More, Peggy                      67-1213 | Orgel, Doris                      65-17782
Oliver, John                      67-92651 | O'More, Peggy                      67-8666 | Orgel, Doris                      66-14336
Oliver, Michel                    66-21480 | Omura, Yuriko                   AC66-10877 | Orgel, Doris                        67-246
```

FIGURE 1.—*Name Index to the MARC Tapes*

authors and their names are far less homogeneous than the limitation to English might suggest. No body of onomastic data including Sadik al-Azm, L'Harl Copeland, Michael De-la-Noy, two persons named D'Souza, Carl C:son van Horn, Klaus aus der Mühlen, Leon Ó Broin, and such a good assortment of M(a)c Gregors, simple and compound, can be considered unsophisticated or trivial. Everhard Ferdinand Im Thurn is misfiled in the index because his prefix was not one of those recognized by the macro instruction used in the program; the omission has been remedied. Appellations such as "A monk of the Eastern Church" proved to be a disaster, but remedial steps here are also plain.

Gaining facility in manipulating machine-readable bibliographical records has been Harvard's chief reward for its participation in the MARC Project. Changing over from an experimental to a production basis is still in the future, and many unanswered questions lie ahead. However, it is hoped that many of the answers will become apparent during the remainder of the work under the National Science Foundation grant. This has been held up awaiting the availability of the MARC II format.

D. Reaction to the Format

It was with some surprise that the Harvard representative early realized that the fixed fields were considered quite controversial by many, since the local reaction was that they were useful and helpful. Fixed fields actually used (in addition to the Library of Congress card number) were the type of main entry, the juvenile indicator, the first fixed field date, the place of publication, the series indicator, and the control indicator. The first language code is to be used in work definitely planned. Length of record was not used in any machine process, but was often used visually in examining tape dumps of problems to be sure that the ostensible and actual lengths were the same (sudden death was near when they were not).

Use of the variable fields ranged from a simple search for the first letter of the prefix of the Library of Congress classification, through the not quite so simple search for any classification number within a certain range of numbers, to a fairly wide variety of word by word and often letter by letter examinations of main entries, titles, collations, subject tracings, and personal and corporate added entries. Personal main and added entries in particular were intensively processed.

The only information added locally was a coded date of receipt inserted in the local information field. Before MARC can be used in production at Harvard (other than in the sense of being a sort of unconventional reference book), local call numbers and subject headings will have to be added. It is expected that the study under the National Science Foundation grant previously mentioned will throw light on the problems involved in these more extensive additions.

Harvard was generally well pleased with the MARC I format, particularly after it learned how to protect itself from the problem represented by the message "trouble counting through record." However, some improvements could be suggested. A search code such as the one now being worked on at Rice University would be a worthwhile addition. The date of creation or initial distribution of the record would be useful in maintaining current and less current files (admittedly, this information can be and has been added locally; but it would seem to be of general utility). The character set does not include a sufficient range of diacritical marks; Polish and Hungarian, for example, are two important languages which cannot now be correctly represented. A distinctive code for umlaut to make possible the insertion for filing purposes of an "e" would be desirable, differentiating it from the diaeresis or other diacritics represented by the same graphic in languages where the insertion of the "e" is not called for. It would be desirable to have an identifier for forenames, which present acute problems in machine processing; this or still another code could be used for generic author statements such as "A monk of the Eastern Church" and for those pseudonyms which should not be treated in the pattern of regular surnames ("Anonymous, M.D."). A distinctive tag for personal names used as subjects is also desirable so that all names are machine identifiable as such, an important requirement when the authority file problem is faced up to.

A somewhat painful topic for which there is no specific location in the outline can perhaps best be brought up here. While a certain number of records defective in a machine sense were issued, especially in the early days of the project, this aspect of error was well within the range to be expected in a pioneering project.

(Three such records, or approximately one in ten thousand, remain in Harvard's current file, and at this point it is admitted to be Harvard's responsibility to remove them.) So far so good. However, the incidence of simple typographical errors has been unhappily high and is very distressing. One of the major virtues of having machine-readable data is that information once brought into correct form can be replicated by machine without the danger of introducing human error. This advantage is lost when incorrect data are introduced into the system. More understandable than the straight typos but no less to be corrected are the fairly large number of misassigned type of main entry codes and misplaced or omitted delimiters. Omission of the delimiters for an author's dates is especially common in author-title, subject, and added entries.

E. The Programs Supplied

The 1401 programs supplied by the contractor suffered severely from a policy decision that they would be written for a stripped down computer with no special features whatever. One wonders how many 8K 1401's, if any, there are in existence that lack some of the more widely available special features such as high-low-equal compare, sense switches, and advanced programming. The avoidance of the latter, including the all-important feature of indexing, made the programs a tour de force in the juggling of addresses, difficult to write, difficult to comprehend, and difficult to modify. Many things that were done very indirectly and deviously could have been done in a simple and straightforward manner, taking far less core, with the use of indexing and the storing of address registers. Apart from this general criticism, there could be many specific criticisms of details of the program, such as the hanging punctuation left at the end of short titles used as tracings in the catalogue card printing program. Comments and section headings in the programs were uneven, there being enough to suggest that they were more thorough than they actually were.

F. Computer Use and Local Programming

The convenient location of the 1401 computer has already been described. Generally speaking it was possible to get time on this machine when it was wanted, or with relatively little delay. Programming for the Harvard aspect of the MARC Project has until very recently been entirely in the hands of the writer of this report. Hence whatever problems existed, they did not include that of a communications gap between librarians and programmers.

Several of the programs written have already been alluded to. All were writen in Autocoder. MCSEL (MARC Selection) reads each MARC record and successively presents first its fixed and then its several variable fields for processing by an *ad hoc* extension. Linkage between the main program and the extension is by means of actual addresses, all ending in even hundreds, specifications for which are given in the program description. At the end of the examination of each field, there is the option to reject the record, select the record, continue the examination, or end the run (with or without selecting the current record). An often used utility extension for MCSEL was that to select records specified by call cards in the same format as that used with the supplied programs. Two programs, MCWBF and MCPBF, to write and print brief entries respectively, represent the first phase of Harvard's efforts to make a better index to the MARC tapes. The first writes a tape with a fixed length index entry for each personal name encountered and for the title, modifying prefix names and removing initial articles but setting codes so that they can be restored. The second prints the sorted output from the first, restoring articles and prefixes to their original status with few exceptions. A later generation of programs produced the index, a sample of which appears here. MCWNI (MARC, write name index) extracts names, stores them in printing form, develops a special filing field, writes a tape of tentative index entries, and prints an edit list of problems. MCENI rewrites the tape produced by the preceding as directed by punched card input and prints a list with forms chosen in one column and those rejected in another. MCPNI prints a triple column upper and lower case listing from a sorted edited tape.

The merge program MCMRG handles all four files as originally sent out (but drops the data from file 2) and adds the date of receipt of new records to the local use field. A modified version, MCMG1, is limited to one file but will handle multi-reel input, i.e., successive reels of one logical file. TWDMP is not strictly speaking a MARC program but has proved indispensable in dealing with MARC. It reads tape in load mode and prints the content, the word marks, and identification of normally non-printing characters, thus giving a complete picture of the tape with character numbering agreeing with that in the descriptions of the MARC format.

A whole family of macro instructions (and macro constants) have been written, of which those most closely related to MARC will be briefly described. RTAPE not only reads tape, with word marks if desired, and has the usual error and overflow checks, but has provision for spacing through a compound fixed and variable length record typified by MARC I. The index register may be freely incremented during the processing of each variable field and will be automatically reset before addition of field length is made to find the next. Late in the year, safety features have been added to prevent the program from hanging up on a record whose field lengths do not add up properly or do not come in the right places. MCHAR is a general character translation and moving routine, used in conjunction with one of several tables. TMARC is the table which provides for translation from MARC codes to those required for the Text 90 print chain, while TMC48 is a table to reduce MARC codes to the 48 character set. Should a different print chain be acquired, or MARC change its character coding, extensive and minute revision of programs would not be required; it would suffice to revise the table macros and reassemble.

The most recent and largest (1.8K) addition to the family of macros is NAMES. This derives a filing field from a personal name, inserting codes for end of surname, end of other name, end of name prior to date, and final end of name including first date (if present) to produce correct filing relationships between simple forenames, simple surnames, compound surnames, and similar names with and without dates. The problem of complex forenames, that is, forenames followed by a comma and various words of description, is not amenable, or at least not readily amenable to this treatment. The macro also closes up the spaces in prefix names, expands M' and Mc to MAC and St. and Ste. to SAINT and SAINTE, and can insert an "e" after an umlauted a, o, or u and an extra "a" when Swedish å is encountered. A few copies of a description of the earlier macros were xeroxed, but this document is now obsolete and an up-to-date description is not yet ready.

G. Administrative Experience

Staffing problems were not severe, always remembering that the project was limited to an experimental effort. A substantially larger systems staff would be required before a comprehensive production system could be introduced. Those members of the staff having a professional concern with data processing naturally took a particularly enthusiastic interest in MARC, as did the top administration of the Library and the heads of several of the larger libraries outside of Widener. The librarian whose collection includes material in the character languages of the Far East is especially keen on finding solutions to the problems of machine encoding of these languages, a question that MARC will have to face sometime in the future. The staffs of the Resources and Acquisitions Department and the Catalogue Department were most cooperative in providing the data requested concerning the dates books were ordered and searched and cards were filed. The Catalogue Department set aside a copy of each Harvard card for current English language books so that comparisons with MARC cards could be made. No antagonism was detected, but it is fair to say that many staff members can not yet see how or when the data will be of direct help to them, and indeed this is a question far from fully answered even in the minds of those most actively concerned with MARC.

MARC has had a highly stimulating effect on local automation plans and budgeting. The work under the grant from the National Science Foundation when it is finally completed is expected to go far toward building a bridge from experimentation to the design of a partial production system. MARC was an important factor in raising the amounts budgeted for computer time. The actual time used, and the costs for it, are summarized below. Unfortunately the breakdown available does not distinguish between 1401 time and small amounts of the much more expensive 7094 time. It should also be mentioned that the rate for 1401 time was $40 an hour during part of 1967:

	Hours	*Cost*
Sept.-Dec. 1966	26.65	$799.50
Jan.-Dec. 1967, Programming & testing	28.76	$973.00
Jan.-Dec. 1967, Production	17.95	$869.50

With a few weeks lead time, it was possible to mount a constantly developing experimental use of MARC without becoming hopelessly buried on the one hand, or on the other of being fully caught up with implementation of new ideas and requests for special tasks, a backlog of which still exists. A much longer time frame would obviously be required to institute a production system, and Harvard experiences until now can contribute little toward an estimate. A general observation may be in order. Basic work on a given task may be completed quite quickly. The original creative phase of a new macro or a new program may be largely concentrated in a few days. The time from the end of this first phase to smooth running is nearly always several times longer than the first phase itself. This may be merely a reflection on the crudity of some of the original work, but enough similar comments have been heard from enough different people to lead the writer to think otherwise.

A problem that will have to be faced and is already being faced to some degree in another context in this Library is the extent to which operations such as keypunching or entry of data via tape typewriter and, in the future, the use of consoles, should be concentrated in a data processing department and how much they should be dispersed among the various departments to whose operations they are related. Experience here so far tends to support a moderate though not an extreme degree of centralization.

H. Special Studies

The promptness study which was the first goal has turned out almost to be a case of *parturient montes, nascetur ridiculus mus.* However, the limited conclusions that may be drawn from it are quite interesting and favorable to MARC. Many thousands of order slips were saved and annotated with various relevant dates of subsequent processing. Data were gathered roughly during the academic year 1966-67. If foresight were as good as hindsight, this effort would have begun later and continued later. Originally there were fairly ambitious plans to match such of these slips as contained no Library of Congress card number with the MARC index tapes by using alphabetical comparisons. Programming for this task has not been undertaken because of the pressure of other work, and the test as first run was limited to those slips which contained a Library of Congress card number as well as all of the dates required (date ordered, date of searching by the Catalogue Department after the books were received, and date of catalogue card filing). Thirteen hundred and forty cards with Library of Congress card number, brief author and title, and the three dates were keypunched and after sorting by the LC card number were matched with the complete cumulated MARC tape. The percentage of matches was disappointingly small, but little can be made of this since the most active period of data collection coincided with the early months of MARC when it was known to be quite incomplete. After eliminating errors, only a small sample of 141 cards remained where there were verified matches with the MARC tape and all dates were present. Intervals between the dates were calculated, and the results, limited though they are, give an excellent rating for MARC's promptness in those cases where an LC card number was available. In one quarter of the cases, the MARC tape was received in Cambridge 64 or more days before catalogue searching. In half the cases, it was received 24 or more days before searching, and in three quarters of the cases it was received no more than 10 days after searching. The complete range in this first sample is from 125 days before until 129 days after.

Since the limitation of the first sample to slips where a Library of Congress card number was already available tended to introduce a bias favorable to the early availability of all LC data (including MARC), a second sample was developed, supplying LC card numbers to slips which had lacked them with the aid of the new index. In this second sample of 47, as might have been expected, MARC did not make so good a showing, one quarter of the records being available seven or more days before catalogue searching, one half by 24 days after catalogue searching, and three quarters by 63 days after searching. The extreme range of relative

MARC receipt dates in the second sample was from 101 days before searching to 177 days after.

If the early situation was indeed as favorable to MARC as some of these fragmentary results suggest and MARC's promptness is maintained and even improved, this will of course be a crucial factor in the eventual acceptance and widespread use of the machine-readable tapes. It was lack of availability of the printed cards at the time and place that they were needed that so long limited the usefulness of that service, especially to the larger research libraries. Bold new steps in the National Program for Acquisitions and Cataloging and the widespread substitution of photography of depository cards or proof sheets for the actual ordering of cards have dramatically improved the total net performance of the card system. MARC thus has much more formidable competition from a rejuvenated conventional system than it would have had a few years ago. It will be interesting to see whether it can do still better on promptness while affording the additional capabilities of machine readability.

Some figures have already been given on the amount and cost of actual computer time used. As is usually true, personal costs were very much greater than computer costs. It is difficult to establish exact fractions of their time which staff members spend on a particular project, but if related work such as that on the macros is included with MARC, the personnel costs combined with travel costs and machine costs have been in the neighborhood of $20,000 a year, even though substantial expenditure directly focused on adapting the MARC input system under the NSF grant has been delayed by waiting for the availability of the MARC II format. This sum is regarded by the Library as a capital investment in experience.

Other than the overall subaccount figures, there has been no systematic recording of processing times. However, the following figures are available. The programs which extract names from the MARC tape and write index entries produce on the order of 400 index entries per minute from a lesser number of MARC records. The selection program runs at about 1,250 titles per minute simply in tape passing and framework time. With a moderately complex extension such as that described concerning India, which includes a good deal of character by character processing, the number of records examined drops to 850 per minute. The earlier merge program which had to pass through File 2 turned out something like 700 titles per minute. These figures may give at least the order of magnitude for processing times with this configuration.

I. Distribution Problems

The tapes arrived very regularly, usually on Monday, sometimes on Tuesday, occasionally even the previous Friday. Until the recent adoption of tape snubbers to keep the tape from coming unwound, at times there would be several feet of loose tape that had to be manually rewound with some care. Some of the shipping boxes and their straps were in very poor condition. However, none of the plastic tape cases, known to be somewhat fragile, was broken, and there was never any loss of information on the tape attributable to shipping problems.

The ideal system of local storage of the MARC data is yet to be worked out; it is hoped that current studies at the Institute of Library Research of the University of California will throw light on this problem. Harvard at present has its entire file cumulated in one series on two reels of tape, the second of which is only about two thirds full, but until recently it kept the four chronological series separate, and it expects to start a new chronological series of its own soon. Only in dire circumstances is it considered worthwhile to pass an average of five-sixths of a tape to find a single title. A change to noncumulative distribution will be perfectly acceptable to Harvard and would save tape passing time. An unfavorable opinion on the present author-title file has already been expressed. It should be mentioned that the list of Library of Congress card numbers from Argonne, with the numbers reading vertically in columns, is a vast improvement over the original list with numbers reading horizontally across the line, a format extremely confusing to use.

The cross-reference files, and authority files in general, are one day going to be very important in a complete machine system. Just how they should be organized and accessed is one of the great as yet unanswered questions.

J. Conclusions

As has already been stated, the actual results to this moment in terms of direct results of a production nature to the Harvard University

Library from the MARC Project have been a limited number of special lists selected from a data base which is still relatively small and limited both in time of publication and language. Viewed in this light, MARC could be considered an expensive experiment. However, it has been shown that it is feasible to receive machine readable bibliographic data from another library, to make selections from it, and to reformat it as desired. Much remains to be done before this new tool can be put to really effective everyday use in an old and complex library which has existing catalogue headings and its own classification to consider. It is hoped that considerable progress in solving these problems may be made in the next two years or so, and that within a few years after that random access devices and consoles in the Catalogue Department may become economically as well as technically feasible and provide a real breakthrough on a very major problem.

The Harvard University Library would encourage the Library of Congress to continue the MARC service in its proposed revised format; to bring the accuracy of the data disseminated up to the level prevailing with the printed cards; and to explore exchanges of data with other countries so that the more difficult portion of large libraries' accessions which is in foreign languages will come to be covered within a reasonable time in the future.

ILLINOIS STATE LIBRARY

Robert E. Hamilton

Coordinator, Technical Services

We have been experimenting with using the bibliographic listing to pre-catalog items for the library, assigning cutter numbers, and such other information as may be needed prior to the receipt of the book. This has been working very well and we feel it will be of great help to use in moving fairly large masses of books through our processing section.

Since the Illinois State Library is a latecomer in the MARC I Project, we are just starting to get programs tested on computers that are available in the Springfield, Illinois area. We do, however, find that from all indications the programs will work very well for us, and we anticipate using the card writer program to develop catalog cards that can be used in a classed catalog which we are now changing from the dictionary catalog.

INDIANA UNIVERSITY

Michael M. Reynolds
Assistant Director of Libraries

Introduction

Indiana University, after two years of study and programming, implemented a library computer-application in January 1966, which accomplishes the ordering of a volume, the encumbering of the order, the accounting and unencumbering of the order upon receipt of the volume and the cataloging of the volume upon shipment to the Regional Campus. The application, while computer oriented for many of the repetitious clerical obligations of technical processing, still requires slow methodical manual searching for cataloging information for each volume being ordered, editing of the cataloging information once it is found, and key-punching and verifying of the edited order and cataloging information into 80-column tabulating cards. However, once a volume has been ordered, the key-punched cards are placed in an orderly file and are selected and used repeatedly upon reorders of the same volume. Thus, the searching, the editing, and the key-punching and verifying are minimal for volumes reordered.

In January 1966, Indiana University applied for participation in the MARC (*MA*chine-*R*eadable *C*ataloging) Pilot Project by the Library of Congress. The MARC Pilot Project is a continuing experiment which is designed to have a single agency, the Library of Congress, provide machine-readable cataloging records on magnetic tapes to the participating institutions. The use of the MARC tapes should reduce significantly the limitations in the current application and enhance the development of future library applications.

It should be noted that Indiana University's initial utilization of the MARC record would be supplementary to its own primary computer application. Accordingly, the record itself would have to be edited to conform to this system. However, it is hoped that with a substantial increase in the MARC data base, dependency upon the MARC record for ordering and processing would increase and the other applications would become feasible.

I. Analysis of Application

A. *Input Data*

1. Type of source documents

Input data for this application will come from two sources, the Library of Congress and the Regional Campus Libraries Department of Indiana University.

a. *The MARC Magnetic Tape File*

Each entry within this file contains cataloging information in excess of that which is now entered into the Regional Campus Libraries application. Usable and desirable cataloging information will be extracted for the RCL (Regional Campus Libraries) application ignoring extraneous cataloging information.

The MARC File should be purged of entries for volumes which are unlikely to be ordered. The responsibility of choosing the entries for purgation would fall on RCL staff, except in the instance of Juvenile Volumes which will be excluded automatically by Data Systems and Services Department. The accuracy and value of these choices would be tested and assessed statistically after a year's operation.

b. *Order Initiating Cards*

A card from the RCL office will initiate the extraction of cataloging information from the "Active" MARC File. The same card will supply unique order information necessary to write the purchase order for the volume. The card will contain the same information as the current RCL reorder card.

c. *Segregation Card*

A card from the RCL office will initiate the selection and retention of the "Active" MARC File entries from the MARC File.

Segregation cards will be prepared twice monthly by the RCL office.

2. Volume of Activity

The volume of activity should be discussed by type of input:

a. *The MARC Magnetic Tape File*

The MARC File will have a rather uniform weekly increase at an arithmetic-progression. The MARC File potentially will grow to enormity; but by using periodical aging techniques, certain items will be omitted from the MARC File. The selection of only those volumes which are likely to be ordered will eliminate approximately three-fourths of the MARC entries.

b. *Order Initiating Card*

Cards which initiate the purchase orders will experience a weekly growth very similar to the net weekly growth of the "Active" MARC File, unless the scope of ordering which the RCL office performs is increased or decreased. Any change in RCL scope must be reflected in any projection.

Considering the fiscal year of a "stable-scoped" RCL operation, the orders placed have a very high weekly activity for one month at the beginning of the fiscal year, a rather uniform weekly activity for eight months, and a relatively light activity for the last three months of the fiscal year. During the last three months of the fiscal year, the RCL office is key-punching and verifying orders which are held until July 1 for new appropriations and are then released in large numbers for about one month.

c. *Segregation Card*

Cards to segregate the MARC File into "Active" and "Inactive" MARC Files will be a rather uniform, semi-monthly input into the MARC application. The size of the monthly segregation card activity will approximate the month's new entries which will be classed as "likely" orders and placed in the MARC File.

3. Frequency of Input Data

Discussion will continue by type of input record.

a. *The MARC Magnetic Tape Files*

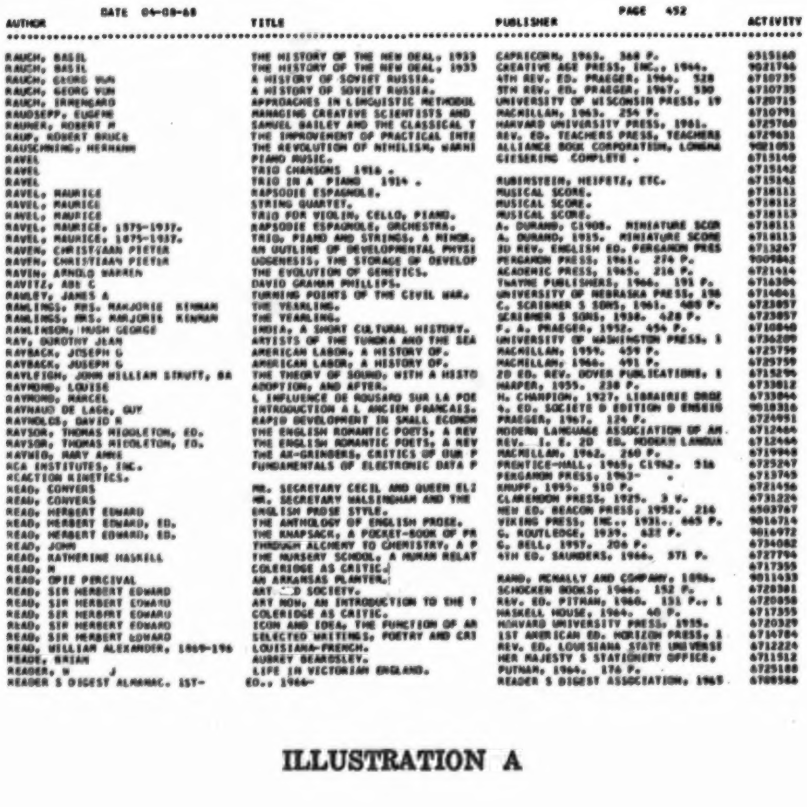
The "Active" MARC File will be "updated" semi-monthly with the addition of selected new MARC File entries and periodically by the deletion of aged entries.

b. *Order Initiating Cards*

The RCL office will forward order-initiating cards semi-monthly to Data Systems and Services Department. The forwarded cards will be processed during the normal weekly schedule of RCL applications.

c. *Segregation Card*

The Segregation cards will be submitted by the RCL office semi-monthly and will be entered into the system at that time to transfer cited MARC entries to the "Ac-



```
                DATE  04-08-68                                                          PAGE  452
AUTHOR                              TITLE                               PUBLISHER                           ACTIVITY
*****************************************************************************************************************
RAUCH, BASIL                        THE HISTORY OF THE NEW DEAL, 1933   CAPRICORN, 1963.  368 P.            6515160
RAUCH, BASIL                        THE HISTORY OF THE NEW DEAL, 1933   CREATIVE AGE PRESS, INC., 1944.     9021746
RAUCH, GEORG VON                    A HISTORY OF SOVIET RUSSIA.         4TH REV. ED. PRAEGER, 1964.  528    6710735
RAUCH, GEORG VON                    A HISTORY OF SOVIET RUSSIA.         5TH REV. ED. PRAEGER, 1967.  530    6710735
RAUCH, IRMENGARD                    APPROACHES IN LINGUISTIC METHODOL   UNIVERSITY OF WISCONSIN PRESS, 19   6720715
RAUDSEPP, EUGENE                    MANAGING CREATIVE SCIENTISTS AND    MACMILLAN, 1963.  254 P.            6710791
RAUNER, ROBERT M                    SAMUEL BAILEY AND THE CLASSICAL T   HARVARD UNIVERSITY PRESS, 1961.     6725760
RAUP, ROBERT BRUCE                  THE IMPROVEMENT OF PRACTICAL INTE   REV. ED. TEACHERS PRESS, TEACHERS   6729631
RAUSCHNING, HERMANN                 THE REVOLUTION OF NIHILISM, WARNI   ALLIANCE BOOK CORPORATION, LONGMA   9021053
RAVEL                               PIANO MUSIC.                        GIESEKING  COMPLETE .               6715140
RAVEL                               TRIO CHANSONS  1916 .                                                   6715142
RAVEL                               TRIO IN A  PIANO   1914 .           RUBINSTEIN, HEIFETZ, ETC.           6715141
RAVEL, MAURICE                      RAPSODIE ESPAGNOLE.                 MUSICAL SCORE.                      6718111
RAVEL, MAURICE                      STRING QUARTET.                     MUSICAL SCORE.                      6718112
RAVEL, MAURICE                      TRIO FOR VIOLIN, CELLO, PIANO.      MUSICAL SCORE.                      6718113
RAVEL, MAURICE, 1875-1937.          RAPSODIE ESPAGNOLE, ORCHESTRA.      A. DURAND, C1908.  MINIATURE SCOR   6718111
RAVEL, MAURICE, 1875-1937.          TRIO, PIANO AND STRINGS, A MINOR.   A. DURAND, 1915.  MINIATURE SCORE   6718113
RAVEN, CHRISTIAAN PIETER            AN OUTLINE OF DEVELOPMENTAL PHYSI   3D REV. ENGLISH ED. PERGAMON PRES   6713267
RAVEN, CHRISTIAAN PIETER            OOGENESIS, THE STORAGE OF DEVELOP   PERGAMON PRESS, 1961.  274 P.       9009862
RAVEN, ARNOLD WARREN                THE EVOLUTION OF GENETICS.          ACADEMIC PRESS, 1965.  216 P.       6721414
RAVITZ, ABE C                       DAVID GRAHAM PHILLIPS.              TWAYNE PUBLISHERS, 1966.  191 P.    6716306
RAWLEY, JAMES A                     TURNING POINTS OF THE CIVIL WAR.    UNIVERSITY OF NEBRASKA PRESS, 196   6714041
RAWLINGS, MRS. MARJORIE  KINNAN     THE YEARLING.                       C. SCRIBNER S SONS, 1961.  400 P.   6723057
RAWLINGS, MRS. MARJORIE  KINNAN     THE YEARLING.                       SCRIBNER S SONS, 1938.  428 P.      6723057
RAWLINSON, HUGH GEORGE              INDIA, A SHORT CULTURAL HISTORY.    F. A. PRAEGER, 1952.  454 P.        6710840
RAY, DOROTHY JEAN                   ARTISTS OF THE TUNDRA AND THE SEA   UNIVERSITY OF WASHINGTON PRESS, 1   6736209
RAYBACK, JOSEPH G                   AMERICAN LABOR, A HISTORY OF.       MACMILLAN, 1959.  459 P.            6725750
RAYBACK, JOSEPH G                   AMERICAN LABOR, A HISTORY OF.       MACMILLAN, 1966.  491 P.            6725759
RAYLEIGH, JOHN WILLIAM STRUTT, BA   THE THEORY OF SOUND, WITH A HISTO   2D ED. REV. DOVER PUBLICATIONS, 1   6715296
RAYMOND, LOUISE                     ADOPTION, AND AFTER.                HARPER, 1955.  238 P.               6733012
RAYMOND, MARCEL                     L INFLUENCE DE RONSARD SUR LA POE   H. CHAMPION, 1927. LIBRAIRIE DROZ   6733044
RAYNAUD DE LAGE, GUY                INTRODUCTION A L ANCIEN FRANCAIS.   4. ED. SOCIETE D EDITION D ENSEIG   9010316
RAYNOLDS, DAVID R                   RAPID DEVELOPMENT IN SMALL ECONOM   PRAEGER, 1967.  124 P.              6724951
RAYSOR, THOMAS MIDDLETON, ED.       THE ENGLISH ROMANTIC POETS, A REV   MODERN LANGUAGE ASSOCIATION OF AM   6712464
RAYSOR, THOMAS MIDDLETON, ED.       THE ENGLISH ROMANTIC POETS, A REV   REV.  I. E. 2D  ED. MODERN LANGUA   6712464
RAYWID, MARY ANNE                   THE AX-GRINDERS, CRITICS OF OUR P   MACMILLAN, 1962.  260 P.            6719948
RCA INSTITUTES, INC.                FUNDAMENTALS OF ELECTRONIC DATA P   PRENTICE-HALL, 1965, C1962.  516    6725247
REACTION KINETICS.                                                      PERGAMON PRESS, 1963-    .          6713745
READ, CONVERS                       MR. SECRETARY CECIL AND QUEEN ELI   KNOPF, 1955.  510 P.                6721456
READ, CONVERS                       MR. SECRETARY WALSINGHAM AND THE    CLARENDON PRESS, 1925.  3 V.        6731224
READ, HERBERT EDWARD                ENGLISH PROSE STYLE.                NEW ED. BEACON PRESS, 1952.  216    6503767
READ, HERBERT EDWARD, ED.           THE ANTHOLOGY OF ENGLISH PROSE.     VIKING PRESS, INC., 1931..  445 P.  9016714
READ, HERBERT EDWARD, ED.           THE KNAPSACK, A POCKET-BOOK OF PR   G. ROUTLEDGE, 1939.  622 P.         9016972
READ, JOHN                          THROUGH ALCHEMY TO CHEMISTRY, A P   G. BELL, 1957.  206 P.              6734082
READ, KATHERINE HASKELL             THE NURSERY SCHOOL, A HUMAN RELAT   4TH ED. SAUNDERS, 1966.  371 P.     6727794
READ, M                             COLERIDGE AS CRITIC.                                                    6717355
READ, OPIE PERCIVAL                 AN ARKANSAS PLANTER.                RAND, MCNALLY AND COMPANY, 1890.    9011433
READ, SIR HERBERT EDWARD            ART AND SOCIETY.                    SCHOCKEN BOOKS, 1966.  152 P.       6728381
READ, SIR HERBERT EDWARD            ART NOW, AN INTRODUCTION TO THE T   REV. ED. PITMAN, 1960.  131 P., 1   6720858
READ, SIR HERBERT EDWARD            COLERIDGE AS CRITIC.                HASKELL HOUSE, 1964.  40 P.         6717355
READ, SIR HERBERT EDWARD            ICON AND IDEA, THE FUNCTION OF AR   HARVARD UNIVERSITY PRESS, 1955.     6720329
READ, SIR HERBERT EDWARD            SELECTED WRITINGS, POETRY AND CRI   1ST AMERICAN ED. HORIZON PRESS, 1   6714784
READ, WILLIAM ALEXANDER, 1869-196   LOUISIANA-FRENCH.                   REV. ED. LOUISIANA STATE UNIVERSI   6712224
READE, BRIAN                        AUBREY BEARDSLEY.                   HER MAJESTY S STATIONERY OFFICE.    6711512
READER, W     J                     LIFE IN VICTORIAN ENGLAND.          PUTNAM, 1964.  176 P.               6725180
READER S DIGEST ALMANAC. 1ST-     ED., 1966-                            READER S DIGEST ASSOCIATION, 1965   6708584
```

ILLUSTRATION A

tive" MARC entries to the "Active" MARC File.

B. *Records Maintained*

The unique RCL record will be maintained in the format of the MARC I record.

C. *Reports Prepared*

One unique report will be prepared by the MARC application. The report will be a "Suggestive Current Awareness Listing" which will be formatted as in Illustration A above but will be further arranged in the Library of Congress classification sequence as in Illustration B below.

ILLUSTRATION B

	Science
Q 335.5 .F4	Feigenbaum, Edward A. Ed. Computers and Thought, A Collection of Articles. McGraw-Hill, 1963.
	Medicine
RC 570 .R6	Robinson, Halbert B. The Mentally Retarded Child, A Psychological Approach. McGraw-Hill, 1965.
	Agriculture—Plant and Animal Industry
ST 41	Fuerst, Elinor V. Fundamentals of Nursing, the Humanities and the Science in Nursing. 3rd. Ed. Lippincott, 1964.

This report will inform the faculty and librarians of recent volumes in the various disciplines (Classifications).

II. Summary of Benefits and Limitations

Using the scant statistical information which is available at the present time, the projected benefits from the implementation of the MARC Pilot Project Application will minimize the searching, editing, keypunching, and verifying in the RCL office for 9000–12000 volumes within the fiscal year in which the application is implemented. Each volume which is ordered using MARC Pilot Project data will reduce the number of searches to one, will eliminate the initial editing completely, and will reduce the number of cards to be keypunched and verified to one instead of the current average of six and one-half cards for those volumes which require cataloging information to be keypunched and verified.

Such reductions of personnel time, equipment and supplies, connote a saving and such will be realized, but the saving will be partially nullified by the increased computer time. However, such reductions will free personnel to attack the ever-increasing, ever-burgeoning number of orders which are being placed through the RCL office. During the fiscal year 1967-1968, the number of volumes ordered will more than double.

Within a relatively short time the MARC Pilot Project will allow other library applications which will be of unestimable value to the faculty and the library staff.

III. Other

MARC tapes were supplied by Indiana to two secondary users, Purdue University and the University of Wisconsin. The MARC record was made available to the Graduate Library School and is being utilized by a graduate student in the preparation of a dissertation.

MONTGOMERY COUNTY PUBLIC SCHOOLS

Richard S. Darling

Director, Department of Instructional Materials

The Montgomery County Public Schools was unable to participate actively in the MARC Project. During the period of the project, reorganization of the county Department of Data Systems prevented the assignment of personnel and machine time for this purpose. Though this problem prevented the accumulation of useful information related to use of the tape record, it did reveal that a library system dependent on a computer controlled elsewhere in the same institution and on personnel not responsible to the library system may have difficulties in using computer services consistently and effectively.

NASSAU LIBRARY SYSTEM

Richard A. Pfefferle
Chief, Technical Services

Our program for producing catalog cards on the Univac 1005 was developed in early 1965 with the cooperation of systems analysts from the Univac Company. At that time, we were unaware of any library actually producing catalog cards by computer so we were really working in the dark. Due to a long delay in the production of a print drum with upper and lower case letters, actual production of catalog cards did not take place until November 1965.

The bibliographic data was either prepared here or taken from Library of Congress proof sheets. The data was key-punched into Univac cards using signals for lower and upper case characters. Two lines of data could be punched into one card, so the number of cards in the master deck varied. Our plan included eventual conversion of this material onto tape so we could search it for needed sets of catalog cards by machine rather than manually as we had to do with punch cards. For this reason, we arbitrarily assigned each deck of master punch cards a distinctive item number.

Our program for the production of catalog cards is rather simple in approach. It does not include fixed or variable fields and contains no codes for any of the components of the data. Thus, we are unable to use our machine-readable information for any retrieval or book catalog production functions. In addition, since we desired a print drum with upper and lower case characters and our print drum could accommodate only 63 characters, we had to sacrifice all but four punctuation characters.

Since we are a public library system doing cataloging and processing for 52 member libraries ordering the same title at many different times, the advantage of being able to produce sets of catalog cards on demand was very important since to sort, assemble and store cards is a very time consuming and expensive operation.

It was our hope with the MARC I Pilot Project that we could use the MARC tapes to obtain bibliographical information in machine-readable form without having to create this information locally. We were aware at the time the pilot project was started that our program was not compatible with the format of MARC I with its fixed and variable fields, but we were assured that the tapes, compatible with the IBM 360 and 1401, could be converted at little cost to tapes compatible with Univac 1005.

When our tape drive units were delivered in late 1966, we investigated this conversion. We discovered that to develop a program to use the MARC tapes on our equipment was quite costly. Added to this was the need for programs separate from our existing catalog card program to use the converted tapes to produce the cataloging information for our catalogers. Since MARC I was a pilot project and originally scheduled to end at the end of June 1967, we felt that the benefits we might receive did not warrant the time and money needed to get the program operating.

Even though we did not have the opportunity to use the tapes, we checked on the timeliness of MARC data by using the "Abbreviated Author/Title List of New and Revised MARC Records." For each of the five tapes issued in June, titles from random pages of the listing were checked against the official catalog and the proof sheet file to ascertain the number of titles cataloged before receipt of the tape and the number of proof sheets available for those titles listed on the tape. In addition, the same titles on the first four tapes were checked against the official catalog on March 1, 1968, to ascertain the number of titles for which we would have had MARC tape data available at time of cataloging. The results are summarized below:

Date	Pages Checked	Titles Checked	Previously Cataloged	Percent	Proof Sheets Available	Percent
June 1	1, 10, 20, 30, 40	71	9	12.7	6	8.4
June 8	2, 11, 21, 31, 41, 51, 61	104	13	12.5	7	6.7
June 15	3, 12, 22, 32, 42	74	5	6.8	16	21.6
June 22	4, 13, 23, 33, 43, 53	109	17	15.6	13	11.9
June 29	5, 14, 24, 34, 44	75	9	12	9	12
Total		433	53	12.2	51	11.5

	Je 1	Je 8	Je 15	Je 22	Total
Total Titles	71	104	74	109	358
No. Cataloged before receipt of tape	9	13	5	17	44
% cataloged before receipt of tape	12.7	12.5	6.8	15.6	12.3
No. cataloged to March 1, 1968 after receipt of tape	17	19	25	23	84
% cataloged to March 1, 1968 after receipt of tape	24	18.3	33.8	21.1	23.5

Of the 358 titles sampled, we cataloged 128, or 35.7 percent, but 44, or 12.3 percent, of the titles were received and cataloged before the MARC record was available. In 1967 we cataloged 18,786 titles of which 14,613 represented adult non-fiction books.

Since we were still not prepared to utilize the MARC II tapes, we decided not to participate in this program. We do have the system designed and hope that in the months ahead we can analyze it to see if it would be feasible and economical for us to revise our present catalog program to make it compatible with MARC II.

NATIONAL AGRICULTURAL LIBRARY

Jeanne M. Holmes
Chief, Division of Catalog and Records

The National Agricultural Library has not yet automated its operations which relate to the MARC program. MARC I tapes were used only for very limited experimentation.

From the weekly printouts of abbreviated author/title entries, the Division of Acquisitions selected new or revised items which appeared to be of interest to the Library. The corresponding card numbers were checked on the numerical list, which was turned over to the Automation staff for printouts of the full cataloging record for selected items. These cataloging records were not used for any purpose other than as a source of information in connection with the Library's participation in the development of the MARC format.

The Library used equipment in the Washington Data Processing Center of the Department of Agriculture. The 1401 programs were run on an IBM S/360 in emulation mode. When emulation was discontinued, the Library had no 360 programs in operation for the Center's IBM 360 using Operating System. Programming staff was not available to make the necessary changes in the Library of Congress 360 programs which used Disc Operating System. Printing of cataloging records stopped in mid-October 1967.

The Library plans to use the MARC II format when cataloging activities are automated and to make as much use as possible of information available on MARC tapes. The very limited experience with MARC I made it clear that criteria must be developed for automatic selection of items of interest from the total record. Manual review of the abbreviated author/title list is time consuming and often unsatisfactory because the list contains insufficient information about the work, especially about its subject content. A start has been made toward listing Library of Congress classification numbers most likely to be effective for retrieving records of interest. When programming time permits, this aspect will be explored further. The possibilities of card production from MARC tapes will also be studied.

The National Agricultural Library, along with other MARC participants and with the National Library of Medicine, will continue to work closely with the Library of Congress in all phases of the MARC program. This program will provide an essential element in the National Agricultural Library's progress toward automation.

REDSTONE SCIENTIFIC INFORMATION CENTER

Jane F. Bentley
Chief, Operations Section, Library Branch

The Redstone Scientific Information Center is located at Redstone Arsenal near Huntsville, Alabama, where it serves some 7800 active patrons from the U.S. Army Missile Command, NASA's George C. Marshall Space Flight Center, other Army activities located on the Arsenal, plus Huntsville area contractors supporting the Army and NASA.

Since its creation in 1961, the Center has been actively engaged in the development and implementation of an integrated system known as ALPHA (Automated Literature Processing, Handling and Analysis) designed to automate the functional service, processing and management aspects of the library operation, with primary emphasis on the latter since they consume so much manpower. Under ALPHA–1, book ordering and receiving, cataloging and circulation, serials handling (including ordering and receiving, holdings records, binding and routing), and patron and language control have been operational for varying periods of time using an off-line, batched process, serial system on IBM 7010 equipment. Since 1967, further development has been suspended on ALPHA–1 and efforts concentrated on ALPHA–2, an on-line random-access system utilizing disk storage and designed to handle the presently operating ALPHA–1 subsystems, plus document control similar to that for books, retrospective searching and selective dissemination of information. Univac 1108 with teleprocessing capabilities and IBM 1050 remote terminals are planned for ALPHA–2.

All computer equipment is maintained and operated on a centralized basis, and MARC programs furnished by the Library of Congress were utilized with IBM 1401 equipment with 8000 characters of core storage and four 7-level tape drives. More recently simple sort and trap routines were programmed and used with an IBM 7010.

Original plans called for the use of MARC data as a scanning tool for acquisitions purposes, for machine comparison to RSIC's already automated on-order file, and for identification and utilization of records for cataloging purposes. Actual catalog card printing was never contemplated, but the integration of the MARC data into the local automated book cataloging process was anticipated and, indeed, still is.

Upon actual receipt and analysis of the MARC tapes by both library and computer oriented personnel, data appeared to be comprehensive and usable, but the percentage of applicable records proved quite small in RCIC's case. Additionally, the existing local arrangements for acquisition of library materials and procurement of LC proof slips and printed cards plus information from the National Union Catalog provided a more practical solution to RSIC's cataloging problems on a day-to-day basis. This is particularly true since the automated book cataloging subsystem is not yet operational, though methods of converting bibliographic data to machine-readable form are essentially complete. As pointed out in a brief performance evaluation study by Mr. Jay L. Cunningham of LC's Information Systems Office (Project MARC Performance Evaluation Study No. 1, 8 Aug 1967) only about 25% of the items cataloged by RSIC appeared on the MARC tapes prior to their processing. For acquisitions activity, less than 1% of the MARC records had not already been ordered and received. The time period covered by Mr. Cunningham's report made it inconclusive but continuing spot checks bear out the trends indicated. Thus MARC has had limited practical utility in RSIC's presently operating manual system, and constraints on manpower and equipment precluded devoting much programming time to experimentation with the tapes. Simple extracts of complete records from the tapes using LC-furnished 1401 programs provided all the data required for acquisitions except cost and the necessary cataloging information. Simple sort and trap routines were devised for the IBM 7010 to provide a complete print-out of the records by LC number and by main entry. The records in this form have proved usable for both acquisitions and cataloging.

Air mailing the magnetic tapes is a satisfactory means of distribution for MARC records, though anything less than the current weekly cycle seems unfeasible. Timeliness is essential in a scientific and technical library such as RSIC where efforts are strenuously bent toward a 48-hour processing cycle. Non-cumulating MARC tapes is quite adequate once local update routines have been established, and this seems the most practical approach in cases such as RSIC where MARC will feed into the local system. The abbreviated author/title listing has been quite useful for preliminary scanning in the current manual system, but arrangement by LC number or a more complete main entry listing seems more desirable once local program support makes additional manipulation of the MARC record a reality. Cross-reference tracing files were not used locally.

The content of the Library of Congress record is more comprehensive than locally required. However, the ease of manipulation, the inclusion of price and the provision for local records makes the proposed MARC II format an excellent communication format. The prospects for standardization give great hope for future interaction between libraries.

Perhaps the greatest benefit to be derived by RSIC from its MARC participation is the awareness of and participation in the development of the viable and comprehensive MARC II format in time for this format's consideration and inclusion in the design of the remaining serials and documents of ALPHA–2, though unfortunately not soon enough for the book portions of the system. Conversion efforts will commence and should be usable by the time the content of the MARC II record makes it more suitable to the acquisitions and cataloging requirements of a scientific and technical information center such as RSIC.

RICE UNIVERSITY

Fred Ruecking
Head, Data Processing

The utilization and design of experimental tests of the MARC data base were under the general supervision of the Advanced Library Systems Project (ALSP) of the Fondren Library, Rice University. Project staff utilized two computer systems during the MARC experiment—a four-tape, 32K-word IBM 7040 and a four-tape, 16K IBM 1401. Use of the MARC data base was intended to be experimental with some operating applications anticipated for restricted purposes. The MARC program at Rice University was aimed toward defining some of the techniques necessary for bibliographic retrieval from bibliographic input. This paper describes the applications, programs, experiences and conclusions of the MARC Pilot Project, with some specific thoughts concerning the continuation of the program with MARC II format.

Applications

The ALSP staff implemented four applications, including the primary application of bibliographic retrieval. These applications have been grouped together as (1) research, (2) acquisition list, (3) Thomas Mann Index and (4) Thesis File.

1. Research

Several programs were written for the IBM 7040 to explore bibliographic retrieval. These programs were:

Name	*Versions*
1. MARC translate & code	1, 2
2. MARC update	1
3. Search Load	1, 2
4. Search-Print	1, 2, 3, 4

Comparable programs were also written for the 1401 when it became available for library use.

We found that use of the Research Computation Center's 7040 system permitted larger core-loads and consequently fewer runs; the 1401 programs were more efficient because character-handling problems were encountered with the 7040. The only timing comparison available is based upon a complete run of the MARC base against some 400 searches which utilized 6½ hours of 7040 time and 6 hours on the 1401. However, the translate run on the 7040 necessitated 9 hours, but the 1401 needed 3 hours. Approximately 220 hours on the 7040 were used in direct association with the MARC project and some 175 hours on the 1401.

Results from the retrieval experiment have led to the derivation of a search-code approach in which bibliographic input data is coded into a specific structure and compared with a precoded data base. The logic framework producing this structure is almost completed and should soon be available in a report. Initial calculations suggest that the search-code approach will correctly identify a given item with an extremely low percentage of erroneous matches.

2. Acquisitions List

A request from the Order Department for a listing of the contents of the MARC base by card number was received, and a program to produce this bibliographic list was generated for the 7040 and 1401. These programs resemble the LC supplied bibliographic listing but go somewhat further since the full title paragraph is provided and any series note is added. These lists are used to verify main entry data erroneously supplied by our faculty when the card number is known.

3. Thomas Mann Index

Extraneous to the MARC project but related to it is the Mann project of the German Department at Rice University. The MARC I format with some modifications was utilized to generate a machine index to the works of Thomas Mann, reflecting biographical data and items of interest to scholars. (See *Computers and The Humanities,* Vol. 1, No. 3, Jan. 1967, pp. 65–71.)

Several programs were written for this project which were adopted for use in the Thesis project. These were:

1. Load data and format
2. Update and correct
3. Retrieval and print

The Mann system is now being altered to a new format requiring new programs for all project phases.

4. Thesis Project

To avoid conflict with current practices in the cataloging department and to create a useful tool for the library, ALSP staff began the creation of a machine-readable data base of the library's thesis collection. Since the public catalog did not contain author, title, added entry or subject cards for this material, several programs were written to produce various kinds of output from the machine-readable file.

Programs written for the 7040 were as follows:

1. Load data and format
2. Update and correct
3. Retrieve
4. Citation print
5. Catalog card print
6. Book catalog print

Results from this application were a qualified success because problems were encountered in devising an acceptable thesaurus of descriptors for the material. Subject headings were not used.

Computer Programs

All applications for the 7040 were written in COBOL with minor read functions in MAP. Because of the use of two machines and two languages (COBOL and Autocoder), some information is available on the efficiency of COBOL for character-handling tasks.

In terms of coding, COBOL is much easier to use than MAP or Autocoder. In addition, the coded program is readily understood not only in terms of individual instructions but also in the processes being performed. COBOL, then, lends itself remarkably well to transfer of programs from one machine configuration to another and also to translation from COBOL to another language (such as Autocoder). However, the effectiveness of the COBOL code is greatly diminished by the relative ineffectiveness of the generated machine code. Indexing, for example, needs three instructions: clear and add, add, and store. The COBOL-generated code is by no means similar to the code produced by either Fortran or MAP for the same set of character-handling processes. An examination of several such processes indicated that Fortran-generated code was about 10% less than COBOL and that MAP was about 25% less. The implications are that COBOL as a character-handling language can be improved upon considerably, but the cost of an inefficient scanning program in the computer may well be overcome by lower programming cost.

The Autocoder applications are very efficient, since Autocoder is so close to the machine-language. Because the 1400 series is character-based, indexing and other functions require much less time. Comparisons between 7040 COBOL and 1401 Autocoder runs of the same job definitely show an advantage on the side of the 1400 series machine—particularly for long, continuous, and batched runs. This advantage becomes more impressive when the respective tape systems are specified. The 7040 uses 729 V units and the 1401, 7330s. Thus, despite the higher speed of the tape drive and the use of greater density tape in the 7040 configuration, the 1401 is able to compete at a distinct advantage in processing speed when character handling is involved.

Experiences

The experience of using MARC I data has been most beneficial in demonstrating at Rice University the techniques, coding, and utility of processing bibliographic data. Nearly all the identified "faults" of MARC I are now corrected in MARC II. Chief among these is the more accurate representation of the field length in the record directory rather than a length which includes itself, the tag, and the data, as found in MARC I.

Several minor problems were encountered but did not affect the speed of developing minor programs nor the processing capability of the tapes. Some of these are listed below (and in some cases refer specifically to the 7-level tapes):

1. The 8-character leader preceding the MARC I record.
2. The use of capitalization indicators in fixed-field data.
3. The use of capitalization indicators in the tag-indicator fields of variable length tag segments.

4. The lack of a signal to mark end of field and record.
5. The use of tape segment lengths as opposed to field length.

The initial experimentation with retrieving bibliographic data using a search code required the use of the MARC I publisher-code in the input data and the data base. It was quickly found that the amount of time required to establish the correct code by manual look up far exceeded the pertinent value of the information in searching.

Local modifications included the addition of three tags to define items of special interest locally. These were:

00 Descriptor
05 Abstract
19 Sorting title

The sorting title tag was subsequently dropped when the second version of the MARC translate program was implemented. The descriptor and abstract tags are still being used for the thesis file.

Before the 1401 was installed in the library, the greatest problem in our local MARC Pilot Project was the lack of computer time. At the time when the greatest number of our experiments were scheduled, the Research Computation Center was also experiencing its heaviest load. As a result it became increasingly difficult to attain access when we needed it. The length of our runs also created problems, and we were usually forced to checkpoint all runs, which permitted the RCC staff to use our jobs as late evening jobs. Consequently, a run submitted on Monday morning at 10 a.m. would not be completed for our inspection until perhaps Thursday or Friday morning, depending on the existing backlog of batched jobs.

No attempt has been made to utilize the programs supplied by the Library of Congress. Several test runs were made of the programs, but (1) the format was not appropriate, (2) the output did not fit the local hardware requirements, or (3) we were unable to alter the program to make it work as we wished.

The means of distribution has not created any problems. Tapes were normally received on Monday and available for return on Thursday. Because of the flexibility of our use of the data, mailing delays had no effect on our projects.

The participation of the Rice University Library in the MARC Pilot Project has had pronounced effect upon the library, its staff, the faculty and the university administration. An initial result was the close cooperation of the Business Office, library staff and university architects in the design of an addition to the library building. Because of library use of a computer, the addition contains the following features:

1. A vertical communication-case providing power and telephone conduit to all levels of the building.
2. Multi-socket outlets to permit the use of portable equipment.
3. A horizontal grid beneath the technical processing area to provide telephone and power lines for future terminals in an on-line system.
4. A machine room totaling 2400 sq. ft. with full environmental control and power supply for an enormous IBM 360/37 configuration or its equivalent.
5. Provision for expanding machine space.

The results of our retrieval experiment have awakened considerable interest among personnel from our own and other institutions. As a consequence, two major proposals for further experimentation and development have been submitted to federal agencies, and one has been funded. Our order department is now most anxious to have access to computer-processing, and a detailed system study is now in progress.

Furthermore, our initial efforts in library automation have generated interest in the faculties of both the sciences and the humanities. Several faculty members have taken the position that the university must back automation activities in the library. Considerable support for the library's work has developed in the university computer committee—the major campus policy-making group on computer applications. Individuals from outside the library have brought the library's activities to the attention of the university administration, and strong support for our efforts and projects now comes from that important source.

Conclusions

Use of the MARC I data base experimentally and operationally has shortened the time-frame originally established for automation development at Rice University. Our best estimate is that three years and considerable sums of money have been saved by our participation in the MARC project.

Because our experience has been experimental and research-oriented, we may have gained more capability than other participants. In addition our approach to participation has permitted an examination of alternatives primarily because we were not committed to an existing operational condition.

Our conclusions are: (1) The data base should not be cumulated. Cumulation should be the responsibility of those who wish to preserve the data base in its entirety, (2) A card number or standard book number sequence seems to be the most generally useful method of organizing the material for distribution, (3) The Library of Congress may find it useful to arrange relationships with some participants for further experimental development of MARC II regardless of the distribution method which finally evolves, (4) The National Serials Data Program should follow the same path as MARC I to MARC II—first an experimental period and then an operating system.

UNIVERSITY OF CALIFORNIA AT LOS ANGELES

Wendell A. Yeatts, Jr.
Project Librarian, Library Systems Staff

I. User Library

The UCLA Library system consists of the University Research Library and 19 branch libraries. The size of the collection is over two and a half million volumes. The UCLA library system maintains close contact with other off-campus libraries, particularly those on other University of California campuses. The UCLA library plans to acquire its own computer, an IBM 360 model 30, during the fiscal year 1969/70 but has found it necessary to rely on the Campus Computing Network during the period of the MARC Pilot Project.

II. The UCLA Campus Computing Network

At the beginning of the MARC Pilot Project, the UCLA Library planned to use the Computing Network's IBM 7094 for processing MARC data and IBM 1401 for printing. (The 1401 could not be used to run MARC programs because it had only 4K storage capacity.) The Library had its own upper/lower case print chain for use on the 1401.

During 1966, the Campus Computing Network replaced the 7094 with an IBM 360 model 75, and the 1401 with a 360 model 20. At that time, the Library requested 9-channel MARC tapes and System/360 program decks from the Library of Congress. Nevertheless, difficulty was encountered: the MARC programs were written in DOS assembler language; the Computing Network's model 75 is being run under OS/360 with a time sharing system. Implementation of the OS/360 time sharing system was not without problems of its own, and these caused a slow down in experimentation with MARC data.

III. Original Plans and Implementation of MARC

At the beginning of the project, the UCLA Library hoped to use MARC data and programs to produce 3x5 catalog cards. But, because UCLA assigns cutter numbers in a different manner than the Library of Congress, the usefulness of the 3x5 card program was somewhat limited. A program has been written (in PL/1) to list selected records from the MARC tapes. The Library also wanted to produce alphabetic lists for use in book selection.

No plans were made to use MARC data in printing of purchase order forms for two reasons: the scope of coverage was so narrow (current English monographs) that MARC data was of small value to the Library, and many books are received on blanket orders and order slips are never made for these items.

IV. Constraints and Operation Problems

Two of the major constraints in the use of MARC data involved getting computer time and programming help. Consequently, insignificant problems sometimes caused unnecessary delays in experimentation with MARC data. For example, an attempt was made to print the contents of a MARC tape using a utility program on an IBM 360 model 20. The program stopped before a single line was printed. Later, it was found that the model 20 will not print data containing upper- and lower-case letters on an all caps train (as larger model computers with "folded mode" processing will). The same program worked perfectly when run on a system that had an upper/lowercase print train. Other problems such as flaws in the operating system on the 360 model 75 did not help in experimentation with MARC.

V. Results Achieved

The PL/1 program mentioned above is the only operational program that currently exists at UCLA for MARC data. It lists selected MARC records in somewhat the same manner as the MARC Bibliographic Listing program. This program is designed to be an aid to cataloging current English monographs.

VI. Reaction to MARC I Format

At the present, the UCLA Library use of MARC data has been confined exclusively to work with the variable field portion of the record. No modifications have been made in the MARC records, but the general consensus of opinion at UCLA seems to favor the new MARC II format.

Some discussion has been made about the possibility of searching the MARC II fixed field indicators for "Government Publication" (position 32) and "Country Code" (positions 18–21) to retrieve and print out government documents published in countries other than the United States. (As a depository library, UCLA has good coverage for U.S. documents.) However, since the implementation of the new MARC format is so imminent, no programs will be written until sample MARC II records are available.

The local use section of the MARC record is not now being used at UCLA, but the Library plans to use this area to hold such information as the UCLA call number (which differs slightly from the LC call number), location on campus, number of copies, whether acquired by gift, exchange, or purchase, date of record, and so forth.

VII. Reaction to MARC Programs

The UCLA Library has encountered great difficulty in making active use of the programs supplied by the MARC Pilot Project. At the beginning of the Project, UCLA received 7-level tapes and 1401 programs. The 7-level tapes could be manipulated to a limited degree by the use of standard utility programs, and they were listed on several occasions to allow the Library to conduct some special studies (mentioned below). The 1401 programs were of no use to UCLA because they were written to be run on an 8K machine.

As time passed, IBM 360's replaced the older generation of computers available for library use; one, a model 20, was of no help in processing MARC tapes because it had no tape drives. A second computer, also a model 20 but with tape drives, was used to print out information from MARC tapes. The remaining computer, an IBM 360 model 75, has now been successfully used to process MARC tapes.

Generally, UCLA reaction to the programs supplied by the MARC Pilot Project has not been particularly favorable, although many of the factors that prevented successful use of the MARC programs were due to incompatible computer configurations, bugs in the model 75 operating system or PL/1 compiler, and the problem of DOS versus OS. MARC programs may have been of more value to UCLA if they had been written in a higher level language such as COBOL or PL/1. However, this cannot be said definitely without further experimentation with MARC programs.

One lesson that has become clear during UCLA's participation in the MARC Project: more time and money should be allotted to increasing the coverage of the MARC data (to include foreign languages and non-current materials, for example) and less effort should be devoted to programming. The UCLA Library has come to the conclusion that full implementation of MARC will come only after it has written individualized programs suited to its own needs and after the scope of MARC coverage is widened.

VIII. Experience in Relation to Computer Facility and Equipment

Most of the problems involving equipment or use of the computing facility have already been stated more or less explicitly. Neither the original computer configuration at UCLA at the start of the MARC Project nor the present configuration matches the two systems for which MARC programs were written (1401 autocoder and 360 DOS assembler language). Because of the newness of the time sharing operating system on UCLA's model 75, little authoritative consulting assistance was available; occasionally one consultant's advice contradicted another's. Job control language presented its problems; logically, a MARC-type record is described as "VB" (Variable length, Blocked), but File 1 of the MARC tape could not be opened until the record specification was changed to "U" (Undefined) in the job control cards. A large number of these and similar problems, which, when taken singly seem relatively minor, have contributed to the inability of the UCLA Library to put MARC data into everyday operational use.

Another minor difference existed between the print chain owned by UCLA and the character set used by MARC: the left bracket prints as an underscore (_), the right as a number sign (#). UCLA computers have "PN" print chains.

IX. Administrative Experience with MARC

During December 1966, an experienced programmer was hired as a full-time library employee to handle both administrative and technical aspects of the MARC Pilot Project. Unfortunately, this person accepted another position in July 1967. Since then, responsibility for the MARC Pilot Project has been divided among several members of the Library Systems Staff and the University's Institute for Library Research. Work with MARC has continued, however, and the list program was written and debugged by personnel at the Institute of Library Research during the latter part of 1967.

In connection with a series of seminars conducted by the UCLA Library Systems Staff to keep library personnel informed on developments in library automation, a seminar was held on January 18, 1968, to familiarize staff members with recent events in the MARC Pilot Project, such as adoption of the MARC II format, the proposed library character set, and the formation of a MARC-users program-sharing group. The program also included a panel discussion of current projects involving MARC and, of course, plans for future use of MARC.

X. Special Studies

Two special studies dealing with MARC data have been run at the UCLA Library. The first was an attempt to measure the promptness with which MARC records arrived compared to other sources of cataloging data. The LC card numbers of 200 depository cards were recorded and checked against the weekly card number listings sent with MARC tapes. The card numbers were collected on March 6, 1967, but no matches occurred between them and LC card numbers on the MARC lists until March 23, showing a lag of approximately two weeks.

The Library Systems Staff hopes to repeat this study to see if MARC records are now being received more promptly than they were in March 1967.

The second study was conducted in December 1967, and approached the comparison of depository cards with MARC records from a different point of view. The weekly author/title list was compared to entries in the depository card file to see if the MARC tape contained any titles that were not on depository cards. Out of a random sample of 56 MARC titles, only four could not be found in either the deposit file or the "in process" file. The four entries on the MARC tape not found elsewhere could have been lost as a result of misfiling in the deposit card or "in process" file.

XI. Distribution Problems

Distribution of the MARC tapes by air mail seems to be adequate. However, distribution between terminals connected by telephone lines would be more desirable if transmission were not too costly and programming not too elaborate. Such a system would allow quicker receipt of MARC data and immediate feedback from MARC users, and libraries subscribing to MARC services could select (and pay for) only those records that were needed. At the present time a MARC user must take all records whether or not he wants them, a factor that may discourage some libraries (particularly small ones) from subscribing to MARC when the tapes are sold, unless some sort of reduced rate is given to small libraries.

In November 1967, the Information Systems Office at LC contacted MARC users and asked how they would react if the alphabetic author/title file (file 2) were discontinued. A poll was made of library systems analysts at all of the University of California campuses and of the secondary users of MARC. Some of the people did not plan to use MARC and they did not care about retaining file 2. All of the others felt file 2 should be retained, even if LC discontinued the author/title listings. Since LC provides author/title lists with the weekly MARC tape, this problem has not become crucial, although some people working with MARC still regret losing the fixed length file 2 records because they are relatively easier to manipulate than the file 1 records.

The present author/title list would be of great value to Catalog Department personnel if it were cumulative and included all records instead of just new and revised entries. (If printing of a cumulative list each week would be too large a task, perhaps a less frequent periodical listing on a demand basis would be more feasible.) Eventually such listings could be taken over entirely by MARC users who could write programs to fit their individual needs and requirements.

Cumulation of MARC records on each reel of tape is a great help to the user library. If possible, it should be continued. If cumulation of MARC records is too costly or time consuming, some other provision for providing cumulated tape files should be made, such as the updating of one or more duplicate tape files that would be available to MARC subscribers when (and if) a complete MARC file were needed.

XII. Secondary Users

UCLA has supplied information, documentation, duplicate program decks or duplicate tapes to eight other libraries or organizations in California. They are the following:

Institute of Library Research
University of California, Los Angeles
Los Angeles, California

Institute of Library Research
University of California, Berkeley
Berkeley, California

Honnold Library
Claremont College
Claremont, California

L.A.C.C. Library
Los Angeles City College
Los Angeles, California

Modesto Junior College Library
Modesto, California

IBM Research Library
Los Gatos, California

Honeywell Corporation
Los Angeles, California

Systems Development Corporation
Santa Monica, California

Of these organizations, two have written programs to process MARC data or have actually integrated MARC data into their own systems. At Claremont College MARC data is being used to aid the conversion of the Honnold Library shelflist into machine readable form.

XIII. Future Use of MARC Data

On Wednesday, November 1, 1967, a meeting was held at the UCLA Library to discuss possible joint technical processing for the southern campus libraries of the University of California. One topic raised concerned the role, if any, MARC would play in a joint technical processing system. Two suggestions were made: first, the Library's depository card file is nearly overflowing its cabinets and catalogers thought they may be able to discontinue filing cards that are duplicated on MARC tapes. Such a plan will be feasible when MARC's coverage is extended to include all items in a category (such as current English monographs) so that extensive checking is not needed to determine if an entry is or is not on a MARC tape. Second, MARC may be helpful in clearing up the backlog of uncataloged and "brief-listed" books that now exists. However, this second plan may not be economically feasible unless MARC can compete effectively with the current UCLA cataloging system which is based on extensive use of the depository card file.

UNIVERSITY OF CHICAGO

Herman H. Fussler
Director

The University of Chicago is a private institution serving, on campus, approximately 6,000 graduate students and 2,500 undergraduate students. The Library at the present time contains approximately 2.6 million volumes and its expenditures were approximately $3 million in 1966/67.

Prior to the availability of MARC I data, the University of Chicago Library had, with support from the National Science Foundation, embarked upon a major effort to develop and test in actual operations a major, integrated, computer-based, bibliographical data handling system. In the development of this system it had been decided to concentrate first upon the development and specifications for handling bibliographic and processing data elements with tagging codes, definitions, and all the other properties of the basic data base to be utilized in the system. This specification for the system and a large amount of the programming to utilize it had been definitely shaped by the winter of 1966/67. The design of the system to some extent overlapped with the development of the MARC system and there were some fundamental differences in approach. The University of Chicago Library system was intended, for example, to handle data for almost all aspects of book processing, whereas the MARC system was developed primarily to handle cataloging data, with a relatively limited capability for other data. Other differences emerged from some unanticipated changes in the MARC I final design.

A very careful examination of the two systems led the staff at the University of Chicago to conclude that a rather substantial program effort would be required to introduce the MARC I data into the University of Chicago Library system on the one hand, and on the other, a modification of the University of Chicago Library system to the MARC I data concepts would not be responsive to many of the design objectives and operational requirements. Since the MARC I data were to be limited to a portion of the currently-published English language materials received by the Library of Congress, and since it was stated that the MARC I format would be treated experimentally and modified in the light of this experience to a substantially revised version, the University of Chicago Library concluded that it would not be sufficiently advantageous to divert the critically-limited programming staff available to an interim adaptation to accommodate the MARC I data. The tapes were made available to other institutions initially through the Argonne National Laboratory, and the Library of Congress was advised of this position in the event it wished to make other assignments.

The University of Chicago Library contribution to MARC I evaluation and MARC II design was based on experience with its own system rather than on utilization of MARC I data. Evaluation and recommendations for data element coverage and design, item record design, input conventions, fixed code lists, and character sets were made during a number of meetings and discussions with the Information Systems Office staff. The MARC II design will, to some extent, come closer to the concepts and specifications of the established University of Chicago data handling system. The University Library will, of course, wish to undertake the programming required to utilize the MARC II data once the specifications are firmly established. We believe it may also be possible to supply the Library of Congress, if desired, with University of Chicago bibliographical data in the MARC II format either by a modification in the Chicago specifications, or by an automatic conversion.

In 1966 and 1967 the University of Chicago project was utilizing an IBM 360/30 computer on a shared-time basis. This was later upgraded to a 360/40. It is planned in 1968 to transfer the project to a time-shared 360/50 computer with OS operating system and permanently allocated partitions dedicated to online library processing. It should be noted that this mode of operation requires programming modifications that are more extensive than required for off-line, batch processing of bibliographic data.

UNIVERSITY OF FLORIDA

John G. Veenstra
Assistant Director

I. Goals

The University of Florida Libraries MARC Committee was occupied with developing a system for meaningful use of the data provided by MARC and the writing of 1401 computer programs to implement such systems.

In the original proposal for the University of Florida's participation in the MARC Pilot Project, the goals were as follows:

a. Produce a selection tool for the newly formed undergraduate collection.
b. Computer generated printing of orders on multiple order (3"×5") forms and accompanying letter order forms.
c. Create a joint holdings tape showing titles ordered and/or held by Florida Atlantic University and the University of Florida.
d. Provide data for pre-cataloging of MARC ordered titles.
e. Provide unit catalog cards for titles purchased and found on the MARC tape.
f. Produce demand bibliographies.

II. Results

Discussions and programming delayed actual experimentation. The major activity at the University of Florida can be seen in the following report by Mrs. Roberta Orcutt of the University of Florida's Acquisitions Department, which gives the results as well as some of the problems faced.

On October 3, 1967, the University of Florida MARC committee met with the Acquisitions Department to plan an experiment in the use of the MARC tape for generating book orders. Earlier in the year programs were written which would generate all the forms and IBM cards presently used by the Acquisitions Department in processing book orders.

The proposed experiment was to test the use of these programs for three months and to discover, if possible, the major strengths and weaknesses of ordering from MARC and to compare these findings with the present automated system.

Limits in funds and personnel time made necessary a fairly modest experiment. Only titles known to be on the MARC tape were to be selected for the experiment. Therefore, as the weekly tape was received by the Acquisitions Department, it was sent directly to the Computing Center where a special selection list was prepared and returned to the Library. Here, selections personnel chose items to be ordered, indicating them directly on the MARC selections list. At the same time they prepared conventional order cards for these books, which were to be used as a control and to be processed by the Library's present ordering system. In this system, IBM cards representing orders are sent through the University's data processing office, generating multiple book order forms and the letter order form. Both sets of orders were then processed and records were kept of the time lapsed between receipt of orders by Acquisitions and the mailing of order forms to the vendors, as well as of the problems encountered along the way. In addition, when the books were finally received and cataloged a simple "have" code was added to the tape but no records of time involved were recorded as this was beyond the scope of the experiment as planned.

Of the 91 items selected, 78 were actually ordered. Of those not ordered, 6 were already on order or received by the Library, 3 could not be ordered due to a programming problem, 2 had insufficient publisher information, 2 were government documents which require special handling and one proved to be an open entry.

During the test, the time lapsed between receipt of the orders by Acquisitions and the actual mailing of order to vendors averaged about 20 days for MARC orders and 17 days for the control orders. Both of these figures are abnormally high since the three months of the experiment included both the Thanksgiving and Christmas holidays. The lowest and high-

136/137

est number of days were 7 and 33 for MARC and 8 and 37 for the control.

The largest blocks of lapsed time for both was that spent at the Computing Center and at Data Processing. The amount of time actually spent in the Acquisitions Department was essentially the same for both. The time gained for MARC through the smaller amount of key punching required was immediately offset by the necessity for typing addresses on some of the orders and key punching business cards for all. Thus, the costs in library personnel time were almost identical.

The details of the experiment will outline more clearly the methods used and the problems encountered. The selection list prepared by the Computing Center consisted of somewhat more information than that in File 2 on each tape but considerably less, and less expensive, than a printout of the complete record. It included LC card number, author, title, publisher, date, and price when available.

The list was circulated promptly to the selections personnel. Their choices were indicated by red initials written directly on the selection list. The cards which they prepared for ordering control copies included all the above information on the selection list except the LC card number. Both the selection list and the control cards were sent to the Acquisitions Department where the control cards were placed with the other regular in-print orders, and the MARC orders were prepared for use with the MARC tape.

The MARC selections were given to one card checker with instructions to check the order file and to search for the price. The control cards were given to another checker with instructions to find the LC card number as well as checking the order file and finding a price. At the completion of checking, the MARC order went to the key punch operator for preparation of cards for the Computing Center, and the control cards were left in the regular in-print order routine to await processing.

For MARC orders the keypunch operator was required to punch only five brief numbers into a card; the LC card number, dealer code, fund code, price and item number. These cards went directly to the Computing Center where the MARC tape and programs generated library order forms, multiple order forms and a bookkeeping card for each selection.

The regular in-print order waited until the next regular order day (in-print orders are processed twice weekly, Wednesdays and Fridays) when they went to the keypunch operator for the preparation of the cards necessary for the regular machine run order. This required that the operator punch into cards, in addition to the information required for MARC orders, all the information required for the finished order such as author, title, publisher and date, order number and date of order. These cards were sent to Data Processing where the same forms and cards were produced.

Several problems immediately presented themselves in using the MARC method. One of the most annoying was the addressing of the library order form to the dealer. A dealer tape had been prepared using codes already established by the Acquisitions Department. A surprisingly large number of selections proved to be the products of small presses not represented on the tape. This meant that finished orders from Computing Center had to be turned over to typists for addressing. In contrast, the regular machine order draws addresses from a punched card which the keypunch operators include with each order. When a new address is required they merely punch a new card.

Another serious problem was encountered when it was found that a MARC generated bookkeeping card could not be used by the Data Processing computer in handling the library's financial computations. This required the library key punch operator to re-punch all bookkeeping cards received from MARC.

In two cases the printout gave us incorrect publishers and in one case a printer rather than a publisher. These, of course, required cancelling and re-processing orders.

Our program is written assuming the LC card numbers will appear on the tape in ascending numerical order. We found that this is not always true. Since the LC card number is used as access to the tape, the tapes must be sorted before use. The Computing Center estimates the cost of sorting at between $50 and $150 per reel.

Occasionally only the place and date of publication are given with no indication of publisher. In these cases the library must simply wait until more complete information can be found.

On the more optimistic side, ordering from the MARC tape does require less keypunching, with less opportunity for error, than the conventional method. Once an item has been found to be on the tape a great deal of accurate information is immediately available without further searching. This includes, in most cases, all the information necessary for ordering except price and, in the case of small presses, the publisher's address.

Separate but related problems were encountered in the attempt to select and order directly from the MARC prepared selection list. It was found that the information was not complete enough for selection. For example, open entries and unusually expensive items were just not apparent from the information on the list. The physical aspects of the list made impossible the judgements usually required in assigning orders to vendors.

In conclusion, it would seem that, since the greatest advantage in using a computer is to perform large, repetitive tasks at great speed, several criteria would have to be met before any efficient and economical use of the MARC tape for ordering books could be made by the library. The quantity of material included on MARC tape would have to be great enough to cover a large part of the library's book ordering and the volume of the ordering itself would have to be quite large. Several of the problems already outlined would have to be solved, such as the problem of the dealer tape. A simple method of determining whether a requested item was on the tape would be essential.

Last, but very important, a better working relationship must be developed between the library and whichever computing organization will handle its MARC operations.

This report is, of necessity, a preliminary one. The experiment has just concluded. The last MARC order has not yet been returned by the computing center and many of the books already ordered have not yet been received. A later and more detailed examination of the findings and a computing of the costs involved will present a more clearly defined evaluation of the value of MARC to the University of Florida Library for ordering books.

III. Other Results

Among the results was the creation of subject oriented lists from the MARC tape. A machine search and printout of all items in the first two tapes for the Library of Congress codes for medicine and for law was made and used by the librarians in these subject areas.

Another project, designed to withdraw from the MARC tape only those entries having bibliography notes, was carried out. The entries were to be arranged by subject. The sample run indicated that this sort of bibliography of bibliographies could be produced on demand and that current awareness lists were feasible. The result also suggested that these lists would be useful supplements to such printed sources as *Bibliographic Index*. An estimate of the cost of computer time ($150 for a single complete run) led to the conclusion that this would be an expensive service.

IV. Future

Preliminary papers have now been written for debugging the acquisition process along with several added features such as machine generated claims for unfilled orders, status reports, acquisition lists, etc.

Papers have also been written for MARC oriented serials processing, budget control, and MARC produced cataloging cards.

Furthermore, in planning for computerized pre-order searching, a survey was made in March 1968, comparing 267 recent English language orders received at the University of Florida Libraries with titles on the MARC tape. The method consisted of checking these orders against the alphabetical printout of those titles on the first four reels of the MARC tape. Of these 267 titles, only 58 (20%) were found on the MARC tapes of which 13 were on reel 1, 12 on reel 2, 10 on reel 3, and 23 on reel 4. This would indicate that, until MARC becomes more inclusive and a large file of titles becomes available, pre-order searching via the MARC tape will not be very productive.

V. MARC Committee

The MARC Committee at the University of Florida has consisted of an administrative member (the chairman), a representative of the reference department, the cataloguing department, the acquisitions department and a programmer from the University's computer center. Each of these have carried a full work load besides their MARC responsibilities. Although the developments have been slow, we have felt that the Committee approach to the development of a system of MARC usage has been fairly satisfactory up to this time.

It has been our experience that the spectrum of points of view, special problems, and special needs of the various areas of activity represented by committee members has made possible development of an integrated system approach to usage of MARC data.

Furthermore, Committee members as staff of their various departments, have been able to communicate MARC activities to their personnel within these departments, with the result that feed-back from departments affected by MARC activities has been readily available, and information relative to the MARC project has been able to be disseminated with ease to many library personnel.

VI. MARC II

In looking forward to future developments, and particularly to the advent of the MARC II tapes as well as switching from the IBM 1401 to the IBM 360, it is necessary to reprogram and to lead to an expanded and more perfect system. For full advantage it is now evident that we need to acquire additional personnel including a systems analyst and programmers, as well as additional equipment more readily accessible to library personnel.

UNIVERSITY OF MISSOURI

Daniel W. Doell
Coordinator, MARC Project

The University of Missouri Library, Columbia, falls into the "large university library" category. The collection consists of upwards of 1,300,000 physical volumes, contained in a main building with eight subject divisions and eight branches. Seventeen separate card catalogs are currently being maintained. The book budget for this fiscal year approaches $700,000.

The data processing equipment utilized in our experiments with the MARC Pilot Project tape included:

- IBM 1440 Data Processing System. This is located in the library building and is used almost exclusively by the library.
- 8K core
  - indexing/store address register feature
  - sense switches
- 1447 Console, with printer
- 1442 Card Read-Punch
  - Select Stacker
- 1443 Printer
  - 52-character type bar (modified H configuration)
- 1012 Paper Tape Punch
  - Tape punch read feature
- 7335 Magnetic Tape Drive
  - ½ inch tape
  - 7 channel
  - 556 BPI
- 1311 Disk Storage Drives (4). 8 million characters simultaneously
  - Direct-Seek feature
  - Scan Disk feature
- FRIDEN flexowriters. Various models were used, all with a special library keyboard, our own specification.

The first objective of the University of Missouri Library in the utilization of the MARC tapes was to prepare catalog cards using MARC data by means of punched paper tape output from our computer that would, in turn, be read by Flexowriters. The Flexowriters would produce the actual catalog cards on continuous stock, which would later be cut apart. The computer would have pre-sorted the data for the various unit cards, shelflist, subject catalog, author/title catalog, etc., so that the Flexowriter output would consist of cards arranged in order and ready to file in the various catalogs.

Our initial efforts were hampered by the fact that the tape drive ordered for our computer was not delivered until March 1967. Therefore, a programming effort was necessary in order to convert the MARC tape records into punched cards which could be read on the library's computer. This process was carried out on the University Data Processing Center's IBM 1401 computer. Programs had then to be developed to read the cards on our computer and store the data on its 1311 disk files for actual use. All of this preliminary effort became useless when the tape drive finally arrived. At this point the programs needed to be rewritten to go directly from tape to disk without the intermediary steps involving punched cards. Next, a program was written to take the MARC records stored on disk and punch a complete card set on paper tape from it. This tape would produce sets of catalog cards when run on our Flexowriters.

The difficulties encountered in writing this program stemmed mainly from the fact that the number of positions that the original record occupies in core has very little to do with the number of characters needed to punch the paper tape which arranges the data correctly on the card when typed by the Flexowriter. One character from the MARC tape frequently has to be punched into the paper tape as a two or three character group; a standard '&' code, for example, would have to be recoded to an upshift code, a numeric 8 code, and a downshift code, in order to type the '&' on our Flexowriters.

These code manipulations not only require careful adjustment of all field length counts, but also make determining the length of the line to be typed on the Flexowriter quite complicated. Some of the codes are non-spacing, such as upshift and downshift codes, and therefore should not be counted when determining the length of the line in the output. The tabulation code, while occupying only one position in the computer, will account for an indefinite number of spaces on the card, depending on where the Flexowriter carriage happens to be at the time the code is read and where the tab stops have been set. The backspace code counts negatively; you have to subtract one from the length of the line for each backspace encountered. All of these manipulations resulted in a program of considerable complexity.

Our decision to use paper tape output from the computer meant that the programs supplied to us by the Library of Congress were of relatively little use, since they were designed to print out catalog cards directly from a computer printer. Only the broadest outline of the procedures involved was of any direct value to us in writing our own programs. The LC programs could not have been used directly in any case, since they were written for an IBM 1401 computer with a printer having a considerably larger character set than the 52 on our 1443 printer.

We do have a suggestion concerning the distribution of the tapes. Cumulating them for an indefinite period would seem to be rather awkward at best, and dangerous at worst. Constant shipping back and forth of large numbers of tapes would be an involved procedure and would result in expensive and frustrating difficulties when the inevitable happens and one or more tapes get lost in the mails. We suggest that the data tapes might be cumulated up to the amount that will fit one reel of tape. A cumulating card number index could be maintained on a separate tape. This index would consist solely of card numbers and reel numbers and would cumulate indefinitely.

A preliminary run of a batch of card numbers against this index tape or tapes would then indicate the data tapes that would be needed. This procedure would eliminate having to search any card numbers against reels in which they are not contained.

Due to the tentative nature of the MARC I format, we did not attempt to write detailed programs to integrate the MARC tapes into our technical processes during the Pilot Project. Plans are formulated, however, for the new system in which these tapes would be utilized when the format has been standardized. When it is in complete operation, the MARC data will be used as a standard source of bibliographic data for all purposes, including the ordering process in those cases when the MARC copy is available before the order is placed.

All original cataloging performed in our library for materials having an available MARC record will be coded so that a complete shelflist on magnetic tape will be maintained in the appropriate format. This shelflist will eventually be used in a more sophisticated system involving a book catalog, information retrieval, etc. For the present it will be maintained in addition to the present card shelflist until appropriate hardware is available to make such applications possible at an economic level which compares favorably with present methods.

We expect that this system should produce an immediate and substantial reduction in the cost and time necessary to prepare catalog cards when contrasted with our current semi-automated procedures using Flexowriters. The necessity of keyboarding a large portion of the data will be eliminated, and the computer will be utilized to expand the basic data record into a complete card set automatically.

Our conclusions from the MARC Pilot Project lead us to favor the plans for distribution of cataloging data in the MARC II format. We intend to exploit the opportunities offered by this service as best we can. The sooner the MARC II format and the distribution method is finalized, the sooner we can begin work on the actual programs needed to implement our system. We therefore urge that these basic decisions be made definite as soon as is practical.

UNIVERSITY OF TORONTO

Ritvars Bregzis
Assistant Librarian

I. Introduction

The University of Toronto/MARC Pilot Project conducted from November 1, 1966, to June 30, 1967, has attempted to explore the derivation of bibliographic data from the LC/MARC machine-readable records and the integration of the derived information into the existing customary processes and bibliographic data files of the University of Toronto Library.

This exploration has involved two major areas of study: the effect of machine-readable data and associated processes upon the customary procedures and administrative aspects of the cataloguing process and the suitability of the LC/MARC format of the machine-readable record for the customary requirements of the Library's existing catalogue records. For purposes of this exploration the pilot project has accepted the existing catalog mechanism and the existing general cataloguing procedures as given conditions. It is believed that it is essential to introduce automation of bibliographic data exchange, derivation, and absorption by gradual shifting of the existing procedures rather than by abrupt change.

Within these constraints the U. of T./MARC Pilot Project has been concerned with procedural arrangements, characteristics of LC/MARC data, compatibility or convertibility of the LC/MARC data versus data originating from other sources, the machine system, the method of communication of the LC/MARC data, and the actual and potential value of the MARC-type system as a service.

II. Development

The described mode of implementation of the University of Toronto MARC Pilot Project was chosen in March 1966 shortly after the University of Toronto Library was informed that it had been selected to participate in the Library of Congress MARC project. The planning of the U. of T./MARC system began immediately, and the general design of this system was completed by the end of April 1966. At that time it was expected that the system would incorporate at least some of the MARC users' programs which were to be developed by the contractor for the Library of Congress in time to be absorbed by the participants for the September 1, 1966, beginning date of the MARC Pilot Project operation.

The LC/MARC operation did not begin until November 1966, and as late as summer of 1966 it was not possible to obtain from the contractor the required specifications of the users' programs which it was developing for the MARC users.

In order to prepare for the use of the LC/MARC records in the planned manner, it was decided in the summer of 1966 to proceed locally with the writing of the programs which we had expected to receive from the contractor. This decision made possible the operational implementation of the U. of T./MARC project in March 1967, which would not have been possible had we allowed the pace of the development of the U. of T./MARC programs to be determined by the delivery of programs from the contractor.

In view of these complications the time allotted to the MARC pilot project, November 25, 1966, to June 30, 1967, by necessity divided itself into two periods: the developmental period until the latter part of March 1967 and the operational period, April 1 to June 30, 1967. The first of these periods was spent in completing and checking out the required programs and in establishing procedures both for the machine system and for the handling of the U. of T./MARC processes in the Catalogue Department of the Library.

The second period was devoted to the operation of the developed system. During the second period only minor adjustments were made in the programs and the procedures. Also an attempt was made to collect the available data concerning this operation.

III. The Operation

The operation of the University of Toronto/MARC pilot system is based on the general pattern of workflow in the functionally related departments of the Library: the Book Selection Department, the Order Department, the Searching Department and the Catalogue Department (see Table I). Current English language monographic publications, which constituted the scope of the MARC project are acquired by the University of Toronto Library mainly on Dealer Selection Order (DSO) basis and are handled by the appropriate units of these departments.

English language DSO materials are received by the Order Department, reviewed by the Book Selection Department, and sent to the Catalogue Department for cataloguing.

The Proofslips Section of the Catalogue Department, which handles all Western language monographic publications with available LC cataloguing information, records the LC card numbers (found on the verso of the title page) of the received books and keyboards them on perforated paper tape (search record input).

The obtained search record is then compared by the computer with the current LC/MARC tape and all corresponding records are *adapted* to the University of Toronto Library's systems requirements and are printed out on individual data sheets. The data sheets are forwarded to the Searching Department.

The Searching Department, which is responsible for gathering all necessary information (for individual book orders and for cataloguing of acquired books), verifies all entries and the title recorded on the data sheet and notes their correspondence with or deviations from the records and entry forms established in the Official Catalogue. With appropriate annotations the data sheets are returned to the Proofslips Section in the Catalogue Department.

The staff of the Proofslips Section reviews the record printout on the data sheet *versus* the book in hand, indicates any modifications or changes necessitated by the individuality of the book in hand or the established pattern of the University of Toronto Library's Official Catalogue, and keyboards the required changes on perforated paper tape. The books are labeled and forwarded to the stack. The information in the paper tape becomes the change record which is used to modify the originally adapted records which now become *derived* records.

The derived records constitute part of the University of Toronto Library's machine-readable file of bibliographic data. However, before these records are deposited in this master file, they are used for producing computer printed catalogue cards for all required catalogues of the University of Toronto Library system.

The computer produced catalogue cards with all the necessary local modifications and adjustments incorporated into their form are printed in the required filing order within each catalogue group. They are then forwarded to the Card Production Section for filing in the appropriate catalogue.

This operation of the U. of T./MARC system is being continued after the terminal data of the pilot project, viz. June 30, 1967, and is expected to be used for processing LC/MARC data until the MARC II format is announced and the necessary program modifications required by the new format can be made.

IV. Results

During the operational period of the University of Toronto/MARC Pilot Project, from April 1 to June 30, 1967, a total of 9,682 new records were added to the LC/MARC tape. During this period 1,541 LC/MARC records were used for the recording of 1,541 English language monographs from a total of 2,845 such monographs received by the University of Toronto Library during the same period (see Table II).

These totals indicate that the records found on the LC/MARC tape as required for cataloguing of currently acquired English language monographs amount to 15.9% of the new records added to the LC/MARC tape covering only 54.2% of the English language monographs received by the University of Toronto Library, while the total of currently acquired English language monographs at the University of Toronto Library constitute only 29.3% of the total of new LC/MARC records. The weekly variation of these percentages ranges from 7.7 to 57, from 31.8 to 82.5, and from 10.6 to 89 percent respectively.

On the whole, the U. of T./MARC Pilot Project has provided timely catalogue information for one-half of the University of Toronto Library's currently acquired English language monographs. The latter constitute not more than a third of the current English language monographs acquired by the Library of Congress, using one sixth of the LC/MARC tape records.

The average number of new records added weekly by the Library of Congress to MARC tapes has been 745; the weekly average of English language publications received through DSO arrangement by the University of Toronto Library was 218, and the average number of records derived weekly from the LC/MARC tape for University of Toronto Library use was 118.

V. Cost

The total direct cost of the University of Toronto/MARC Pilot Project was $25,108, accounting for salaries of University of Toronto Library staff members who participated in the development of the project, for direct cost of IBM 1401 computer time, and for supplies.

The sum of $16,800 was expended for salaries covering 1 1/6 programmers for eight months,* a half-time operator for the same period, and a total of 238 man-days of personnel from the Library's Technical Services departments. A total of 202 hours of IBM 1401 computer time was bought for the project at $35/hour, and an additional four hours of IBM 7094 time and 10 hours of PDP8 time were used at no direct cost.

The $25,108 represents the direct cost of the project development as well as its operation from April 1 to June 30, 1967. It is estimated that monthly operation costs of the U. of T./MARC project at its June 1967 volume might be approximately $1,400. This would indicate a development cost of approximately $9,000, most of which was expended for programming.

VI. Conclusions

The University of Toronto/MARC system was operated as a pilot operation for three months. Catalogue cards were produced weekly for currently acquired materials and filed into the University of Toronto Library catalogues. The system has already provided a reasonable alternative method for creating catalogue records that can be conveniently integrated with the existing University of Toronto Library catalogue files.

The principal observation of this pilot operation indicates the feasibility of practically acceptable integration of the machine-readable record with the customary method of creating catalogue records. Although there are a number of aspects to be refined before such a machine-oriented technique could be instituted on a large-scale operational basis, the fundamental concept of supplying machine-readable information for local adaption and use can be realized successfully. The service aspects requiring further elaboration include the method of communicating the machine-readable data to the interested party (presently via airmailed magnetic tape reel at weekly intervals), the communication and administration of the authority data for the bibliographic records communicated (for the name and title entries, for the topical terminology, and for the definition of classificatory relationships), and last but not least the promptness of creation and communication of the LC originated bibliographic data.

The duration of the LC/MARC Pilot Project and of the U. of T./MARC project was too short to permit definite conclusions to be drawn with respect to the problems involving the scope of coverage available from LC/MARC. Only quantitative statistics in static relationships are available from the U. of T./MARC Pilot Project. It is hoped that some dynamic aspects on a sample basis will be statistically documented during a limited period in the winter of 1967-68.

The available statistics appear to indicate that approximately one-half of the required bibliographic information arrives too late for use in cataloguing of current materials, if it is included in the LC/MARC tapes at all. It is noted, moreover, that the total of University of Toronto current English language acquistions amount only to one-third of the number of current English language publications registered on the LC/MARC tapes. The possibility that a substantial proportion of University of Toronto Library acquired current English language publications would be exclusive of those acquired by LC appears to be small. The most likely explanation of the quantitative results of the U. of T./MARC project is that there exists considerable adverse timing in the operation of the various functions of the acquisition-cataloguing-communication cycles between the Library of Congress and the University of Toronto Library.

---

* Time spent for program design was done concurrently with another project and is not included.

Both libraries acquire current English language publications on standing order basis and receive these materials likely with only a slight difference in time. This must result in a demand for catalogue copy information on the LC/MARC tape soon after the receipt of the respective materials at the Library of Congress.

The time span of "soon," however, does involve at least one week which is taken at the University of Toronto Library to check in and review the received books and to forward them to the Catalogue Department for processing.

This one week, however, does not appear to constitute the critical time factor in its entirety. The distribution of the required records found on the LC/MARC tape over a period of 11 weekly cycles shows a heavy (80.9%) concentration in the cycle in which the need for the record originated (see Table III). The remainder was found in the four following cycles. The total distribution of the searches for a total of 2,079 required records from a total of 7,417 LC newly added records was as follows:

first try	735	(35.3%)
second try	69	(3.3%)
third try	58	(2.8%)
fourth try	36	(1.7%)
fifth try	8	(0.4%)
Total found	910	(43.5%)
Not found at all	1,169	(56.5%)

Thus more than one half of the required records have not been available from the LC/MARC tape for more than five weeks after the need for these records originated.

In view of this low search yield and the insufficiently explained reasons behind it, it would be desirable to follow through the progress of the entire sequence of processes both in the Library of Congress and the University of Toronto Library for a meaningful and unbiased sample of the same titles, beginning with the book mail opening and ending with mail date stamps on the magnetic tape container.

The procedural arrangements necessary to integrate the MARC derived records in the University of Toronto Library catalogue system appear to be simple, and they follow the basic pattern of arrangements for other materials with available LC cataloguing information. All human operations required for U. of T./MARC follow existing patterns and most are performed by library assistants. The only new operation for the cataloguing staff is the required operation of the paper-tape typewriter, and the only departure from accustomed editing procedures is the review of the fixed field information and the redefinition of the type of entry, itself not normally a difficult task. The principal procedural difference of U. of T./MARC is the bypassing of all card production procedures, and this is a most welcome one. The entire cataloguing cycle for the U. of T./MARC processed materials is settling down to less than two weeks compared with the current average of two weeks for the customary card production cycle alone. On the whole it appears that utilization of machine-readable bibliographic information can be integrated with an existing cataloguing system without disturbing the established processes.

The characteristics of the expression of LC/MARC data constitute one of the principal objects of exploration for the LC/MARC pilot project. From the point of view of the University of Toronto Library as one of the participants in the LC/MARC project these characteristics likewise have been an object of intense interest.

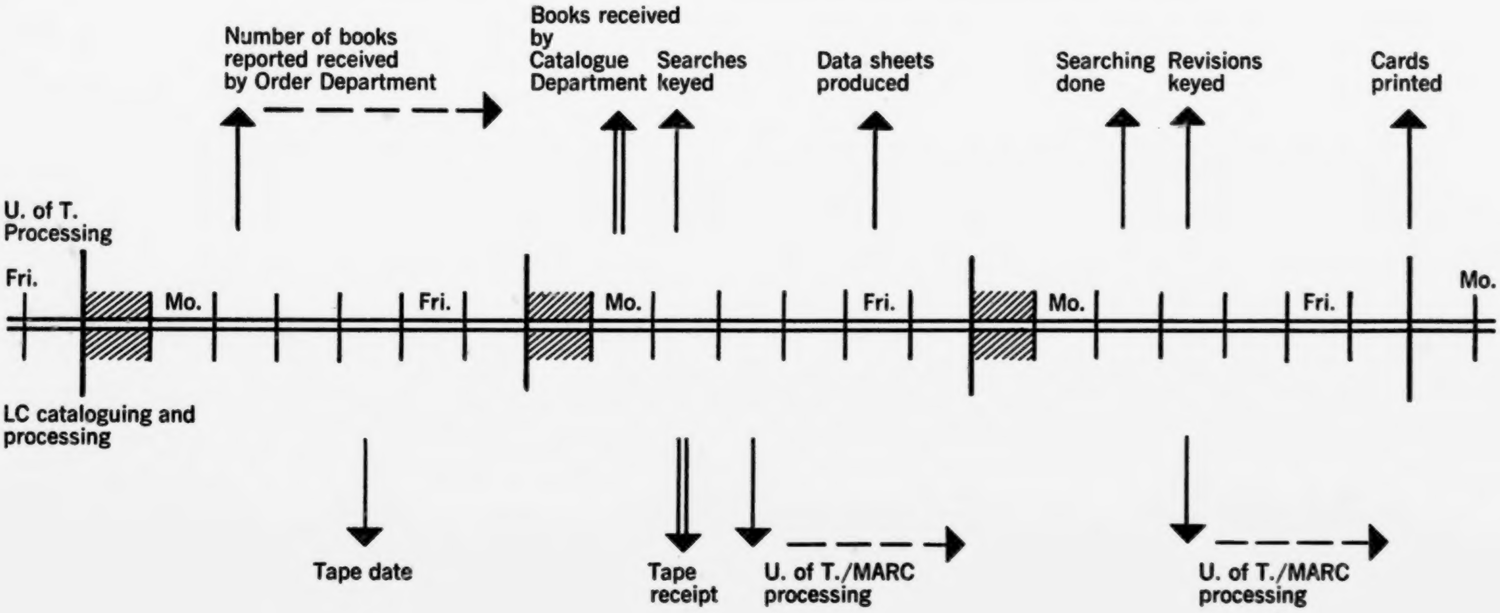
The evaluation of the LC/MARC method of bibliographic data expression, i.e., the MARC I data format, has been explored not only with respect to integration in the existing system of the University of Toronto Library catalogue record production but has also been examined in relation to the requirements of a conversational on-line bibliographic information control pilot system.

From this combined point of view four areas meriting special attention appear to exist:

1. more specific definition and identification of explicit bibliographic data elements and reduced use of coded information in fixed fields;
2. limitation of coded information as much as possible to types of implicit bibliographic information than can be expressed in highly systematized but simple coded patterns;
3. more liberal use of identification points (delimiters) within bibliographic data categories comprised of multiple data elements;
4. a comprehensive system of organization of all bibliographic access data categories, independent of the "entry" system but compatible with it.

Specifically, data categories such as place of publication, publisher, names and titles in contents notes, and other similar data categories merit specific definitions and identification in their own right. Coding of information such as the name of the publisher, on the other hand, may only have a limited practical value. Identification points (delimiters) within the title group separating significant parts of complex titles, names and elements of edition statement within the collation statement and within the contents note are important, particularly in view of future development. And last but not least, the definition and identification of access data categories, i.e., entries and related data, should be expressed in terms of these data rather than entry functions assigned arbitrar-

TABLE I.—*University of Toronto/MARC Processing Cycle*


Number of books reported received by Order Department
Books received by Catalogue Department
Searches keyed
Data sheets produced
Searching done
Revisions keyed
Cards printed
U. of T. Processing
Fri.
Mo.
Fri.
Mo.
Fri.
Mo.
Fri.
Mo.
LC cataloguing and processing
Tape date
Tape receipt
U. of T./MARC processing
U. of T./MARC processing

A typical processing cycle during the pilot period, April 1, 1967 to June 30, 1967.

TABLE II.—*University of Toronto/MARC Pilot Project UTL Derived MARC Records 1966/67*

MARC Tape Date	Date Rec'd at UTL	Cycle	Number of Records on LC tape		English Language DSO Materials Received	Percentage of		
			New Weekly Additions	Derived by UTL/MARC		Derived Records from DSO Receipts	Derived Records from LC's New Additions	DSO Receipts from LC's New Additions
April 6	Apr. 11	15	590	164	[118][1]	139.0	27.8	20.0
13	18	16	875	168	[93][1]	181.0	19.2	10.6
20	26	17	612	186	[545][2]	34.1	30.3	89.0
27	May 2	18	930	104	[327][2]	31.8	11.2	35.2
May 4	12	19	591	70	[144][2]	48.6	11.8	24.4
11	16	20	808	62	[137][2]	45.1	7.7	17.0
18	25	21	220	125	292	42.8	57.0	132.0
25	29	22	1097	127	246	51.5	11.6	22.4
June 1	June 8	23	603	101	172	58.8	16.7	27.3
8	12	24	999	118	334	35.2	11.8	33.4
15	21	25	711	105	169	62.4	14.8	23.8
22	27	26	963	110	133	82.5	11.4	13.8
29	July 5	27	683	101	135	75.0	14.8	19.8
Total			9,682	1,541	2,845	54.2	15.9	29.3

[1] Proportional estimate. In addition a number of delayed receipts were processed during this week.

[2] Proportional estimate.

ily to these data. A name or a title as an access point to an unknown bibliographic record may in the long run be a more likely approach than this name or title can be if arbitrarily defined as only one of several possible types of entry.

It is hoped that the LC/MARC II format will include provision for these and other adjustments necessary to exchange machine-readable bibliographic information in as effective a way as possible.

TABLE III.—*University of Toronto Derivation of LC/MARC Records Distribution by Search Cycle*

Tape Date	Cycle	New LC Additions	DSO English Language Record	DSO English Monographs							
				Derived Total	Found on Search Cycle					Not Found	
					Current	2d	3rd	4th	5th		
June 6/67	23	603	172	91	71	5	9	6	0	81	U. of T./ MARC Pilot Project
8	24	999	334	111	90	7	8	5	0	223	
15	25	711	169	90	52	22	7	6	3	79	
22	26	963	133	88	43	12	24	6	0	45	
29	27	683	135	75	62	5	2	6	0	60	
July 13	28	730	287	151	144	0	0	0	0	136	
20	29	531	251	94	86	6	5	1	1	157	
28	30	424	158	57	54	5	1	1	2	101	
Aug. 3	31	574	103	40	33	2	1	4	1	63	
10	32	643	243	68	57	4	0	1	1	175	
17	33	556	94	45	43	1	1	0	0	49	
Total		7,417	2,079	910	735	69	58	36	8	1,169	
	%			100	80.9	7.6	6.4	4	1.1	--------	
			100	43.5	35.3	3.3	2.8	1.7	0.4	56.5	

WASHINGTON STATE LIBRARY

Josephine S. Pulsifer
Chief of Technical Services

I. Description of User Library

Washington State Library has two prime responsibilities: 1) library service to the state legislature and to state agencies from main and branch libraries, and 2) development of library service throughout the state.

Within the second area of responsibility, three regional library demonstration programs have been conducted. Book catalogs were developed for two of the three demonstration systems.

The Columbia River Regional Library Demonstration, subsequently named North Central Regional Library after its acceptance by the voters in 1960, comprised five counties with relatively few previously existing libraries. A catalog of current acquisitions was developed first, and existing collections incorporated later. The usefulness of a book catalog in serving the present system of twenty-five branches and three bookmobiles, in which collections are constantly shifted, is readily apparent.

Centralized ordering, cataloging, and preparation services are provided at the headquarters library in Wenatchee. EAM and EDP services for book catalog production were provided by contract through 1965.

The Timberland Library Demonstration, still in progress, undertook to provide a union book catalog as a major feature of the Demonstration. Five counties and three municipal libraries in western Washington make up the group of participants. Of the five counties, four had established county library service prior to the Demonstration.

To provide the union catalog, the personnel of the Demonstration Processing Center, located at the State Library, began in 1963 to photocopy the shelflists of the participating libraries. These were combined into a single file in title order, coded numerically for author and title sequencing, edited and keypunched. A local service bureau took it from there through the computer-printed offset master.

Centralized book ordering, cataloging, processing, and production of a book catalog of current acquisitions were undertaken in January 1965, after most of the work on the base author and title catalogs was complete.

During 1965 the State Library's technical services supervisor assumed direction of the Processing Center, and new work flow patterns and a new catalog format were recommended by the consulting of Mr. Joseph Becker and Dr. Robert M. Hayes. After a programmer was added to the staff, the Processing Center staff took over active direction of the Timberland book catalog production, terminating the services provided by the local service bureau. Keypunching, verifying, and offset printing continue to be in-house operations; the remaining EAM and EDP operations are performed on equipment in other state agencies on instructions from the library's staff programmer. During 1966 a computer-produced adult subject catalog was added to the previously issued author and title union catalogs for the region.

The State Library was an eager applicant for participation in the MARC Project. Its book catalog production activities were viewed not as simple production services to the Demonstration (and to North Central Regional Library, whose catalog production was assumed by the Processing Center in 1966) but as a developmental activity toward a possible statewide book catalog in which interest had been expressed by the Washington Library Association.

A WLA Book Catalog Committee had recommended in 1965 that the most useful book catalog would be one representing the holdings of regional libraries, including King County Library, and that the State Library would be the proper agency for the study, development and maintenance of this project.[1]

King County Library System, which today has forty branches serving the county surrounding Seattle, first distributed book catalogs produced with IBM unit record equipment to six branches in August 1951. Author, title, and subject listings of the current collection of each branch are revised every two months. The Library has been interested for some time in developing a union book catalog of all branches, and their union card catalog at headquarters has been extensively reviewed toward this end.

---

[1] Final Report of the WLA Book Catalog Committee, 28 July 1965.

II. Original Plans for the Use of MARC

The State Library saw in the MARC Project a possible solution to the problems of providing book catalogs that would be adequate in the amount of bibliographic information given and adequate in frequency of updating, yet economically feasible for public library systems or groups of systems. It seemed obvious that utilizing cataloging data converted centrally, once and for all, to machine-readable form would be more economical than converting even incomplete data locally.

The book catalog which the State Library proposed as its primary effort in the MARC Project is briefly described as follows:[2]

A. Coverage. Current acquisitions of King County, North Central and Timberland Library systems listed in a combined adult/juvenile catalog.

B. Form.

1. Register catalog presenting all information programmed to print on a 3 × 5 card, modified for page rather than card format, with a local register number added. This catalog would be printed in register, i.e., accession, number order, so updating is unnecessary.
2. Finding list information selected from the total record for separate author, title, and subject catalogs. Information included under each entry: register number, author, short title, publisher, date, holding library systems showing class number and other necessary designations for each system. The register number allows reference to the complete bibliographic citation in the register when required. Access points include all LC tracings, a straight title entry whether traced or not, and additional subject entries where desired. Sorting is by straight computer sort. A look-up table allows the initial English article to be ignored in sorting. Finding lists would be printed monthly on the NUC cumulation pattern.

C. Current processing aids. Multiple main entry cards (for headquarters catalog and/or branch shelflists), punched locator cards (for locating individual copies within a system) and book preparation labels. Basic to the system is the theory that the tape record should be used without having to inspect and edit the record first. When not provided in the MARC record, the Decimal Classification number is added to the subject catalog. Neither becomes a part of the register, which reflects LC information only.

D. Substitution for catalogs in production. It was hoped that MARC coverage would be broad and prompt enough so that a moderate amount of original input of non-MARC titles would allow discontinuance of North Central and Timberland catalogs, provided continued availability of MARC tapes was assured.

Applying the MARC record to the above uses appeared to give the State Library a unique aspect on which to report to the Library of Congress. No other state library was selected among the original participants, and no other participant selected planned to work immediately on a book catalog.

Secondary plans for utilization of the MARC records included the production of catalog card sets for the State Library's catalog and the selection and printing of special bibliographies.

The Washington State University Library at Pullman and the Washington State Department of Public Instruction were designated as secondary participants, but desired only occasional tapes on request.

III. Computer Configuration Used

In mid-1966 a Data Processing Service Center was established within the Washington State Department of General Administration to serve the state agencies who do not require their own computer installation. Available computer facilities, pending installation of a third generation computer in July 1968, consist of an IBM 1401 with the following configuration:

14K core storage
4 7330 tape drives (7 level)

---

[2] Detailed plans for utilizing the MARC record were described in "Washington State Library, MARC Participation Plan," December, 1966.

2 1311 disks
1 1403 printer (48 character print chain)
1 1402 reader punch

After programming for the IBM 1401 had begun, the Library was invited to use an IBM 360 Model 30 in the State Department of Institutions. This offered the advantage of an expanded print chain, providing many punctuation characters available in the 1401 system. Since indefinite availability of the 360 could not be guaranteed, 1401 programs were used throughout the experiment. In October 1967, the Model 30 was replaced by a Model 40. The configuration of the two 360 Models are as follows:

Core storage	64K	128 K
Tape drives	4 729 (7 level)	4 2400 (9 level)
Disks	2 1301	4 2311
Printer	1 1403 (60 character print chain)	1 1403 (60 character print chain)
Reader punch	1 1402	1 1402

IV. Actual Implementation Accomplished

A. Constraints. Availability of titles on MARC tape was the most critical factor in the system. Adoption of the MARC catalog as a substitute for the two operational catalogs would only have been possible if a large enough proportion of the titles currently received for cataloging had been covered. Available staff would not permit original input of large numbers of titles into the MARC format. A high proportion of the titles selected for acquisition by the three public library systems are current American trade publications, listed in ***Publisher's Weekly.*** Since this was the coverage proposed for the MARC Project, it was hoped that a large portion of titles acquired could actually be cataloged from the MARC record. As shown in Figure 4, of 6552 titles cataloged for the Timberland Demonstration during the calendar year 1967, only 2550, or 39% appeared on the tape at any time during the year. Other studies of MARC coverage of titles wanted and availability of the MARC record with respect to receipt of the book and proof slips are reported in Figures 2, 3, and 5.

B. Programming and operating problems experienced. By January 1967 it appeared that programming would take longer than we thought, and it was decided to postpone inclusion of series entries and title *added* entries, i.e., other than the straight title, until a later phase.

Delays in completing work on the Register catalog and in producing multiple main entry cards were the result of 1) delay in availability of the 3×5 card print program and 2) difficulties in adapting the 3×5 card program. The complexity of the 1401 program, our own programming inexperience, and lack of detailed familiarity with the program on the part of the LC staff all contributed to the difficulty of adapting the program to our uses.

Other difficulties and delays resulted from the necessity of changing the print configuration from that used by LC to that required for the 1401 standard print chain and then for the expanded 60 character print chain of the 360. A few small conversion bugs remain. The print configuration change results in an occasional BOOWH on line 2 of a continuation card, between the title and the date. And although the constants in the LC program were changed so that the parentheses and other special characters generated in that program would match ours, one parenthesis is so carefully hidden that we have never been able to find it to make the change.

C. Results achieved. By June 1967 a small sample catalog was issued, but left much to be desired. The register number was faked rather than computer generated; first and continuation cards printed separately rather than a continuous single entry; line spacing had not been adjusted from card to book catalog format; such bugs as cut-off titles were present; the subject catalog was still in process of being programmed.

By October 1967 a second sample catalog of 342 titles, with the above deficiencies eliminated, was issued and widely distributed to Washington libraries. (See Figures 7, 8, 9, and 10).

An annual catalog covering titles acquired by any or all of the three library systems during 1967 and appearing on the MARC tapes from the beginning through December 1967, is now in preparation. This catalog will allow, for the first time, a reasonably accurate analysis of the degree of overlap in the acquisition of titles in the three systems. An effort is being made to include a sampling of descriptive and subject cross-references, as well as added subject headings, in the annual catalog.

3×5 catalog sets, with overprinted headings, and single and multiple main entry cards are now regularly produced. A set of four labels (for book-cards, pocket, and spine) is produced for each copy of a book in two of the three systems. These cards and labels have been incorporated into the regular processing procedures of the library systems and/or the State Library and are used wherever available *on time*. (See Figures 11 and 12).

Machine-readable cards for the central locator file of the Timberland Library Demonstration have been produced, but not on a regular basis to date, for they essentially duplicate the locator cards printed as part of the Timberland book catalog system. The MARC locator cards are computer punched on preprinted cards, which are then interpreted for manual file searching.

V. Reaction to MARC I Format

A. Uses made of fixed fields:

1. LC Card Number (Field 3) used for selecting wanted records from the tape.
2. Date 1 (Field 16) used as source of data for spine and book card labels.
3. Type of main entry, personal or other (Field 5) used to determine the length of entry and the punctuation used on labels.
4. Length of record (Field 25) used in writing our holdings tape from the MARC tapes. It determines where the end of record indicator is placed to stop transmission of information when writing the holdings tape.

B. Special uses made of variable fields. Author, title field up to delimiter, publisher and date portions of imprint field, subject and added author tracings, LC and Decimal Classification number are selected for author, title, and subject catalog tapes. Another selection from author and title fields is made for label production. Decimal Classification and subject tracings were also used as primary and secondary criteria for selecting titles for special subject bibliographies (program incomplete).

C. Local-use data elements employed.

1. Register number: an "accession" number assigned sequentially to each new title entered into the system.
2. Age level: A(dult), Y(oung people), and J(uvenile).
3. Special designations: B(iography), E(asy book), R(eference).
4. Holding library system, and individual holding library in the Timberland system.
5. Number of volumes and copies within a given system.

D. Modifications made locally to format. A 200-character field for local information was added to the LC record.

E. Changes desired in format, with justification.

1. Juvenile indicator:
There is no device for selecting only juvenile titles. Fixed Field 11 contains an X if the form of the subject heading or the classification number indicates juvenile. But, in addition, an X is given in this field for all AC card numbers, which may include young adult titles and books about juvenile literature. This makes it impossible for us to use this field as a juvenile indicator in the book catalog. We request a fixed field indicator for titles with LC or DC juvenile classification or "Juvenile literature" subject subdivisions. Even more desirable would be the application of consistent criteria in designating juvenile literature.
2. Biography
The classification of adult biography with the subject in the 17th edition of the Decimal Classification deals a mortal blow to the biography collection preferred by many libraries, unless the assigned classification number for biographical materials is to be modified. It would be possible to search out the (B) designation which follows the subject classification for some biographical materials, but biography is not very broadly interpreted. Juvenile biography is still classified as J92 by the 9th abridged edition. Autobiography has a special problem in that it requires that a subject heading be added for the person wherever a divided catalog is in use. Collective biography needs a special designation so that its subject classification can be changed to conform to local practice, B being usually reserved for individual biography. To help resolve these problems, we request that a fixed field be provided to indicate individual and collective biography as well as autobiography.
3. We request that provisions be made in the case of a double imprint (2 places, 2 publishers) for any desired selection of items. In the case of London, Batsford, New York, Harper, we want to drop both places and retain both publishers. We should also have the

option of dropping everything up to the second publisher.

4. We request delimiters between parts of the collation, so that paging can be used without illustration and size.

5. We request that the date of first entry of a record on the tape be included in the tape record. This would be useful in analyzing relationship of time of acquisition, date of availability of the MARC record, and date of availability of other sources of cataloging information.

VI. Reaction to LC-Supplied MARC Participant Programs

A. 3×5 card program. As noted earlier, the delay in availability of the program delayed our adaptation of the program to print the register catalog and to produce multiple main entry cards. Also as noted, the program proved very difficult to understand and consequently to change, and little help was available from the Library of Congress, since their staff personnel neither wrote nor used the program.

Beyond the few conversion bugs noted, actual production of card sets was accomplished fairly readily, but no effort was made to make adaptation to the program for printing card sets. We resolved to be content with typing call numbers on the cards until MARC II is implemented.

The format of the 3×5 cards was unsatisfactory in the amount of available space left for a call number. It is understood that an effort was made to get as much text as possible onto the card, but the margin provided required division of the classification number into more segments than is desirable if there is another solution. Using the Decimal Classification, with a maximum length of 8 characters (4 places after the decimal), two lines would be required for the class number, plus one or two lines for the Cutter number. Given the fact that computer print takes far more space than the type-set LC printed card, it seems desirable to use all possible space on the card for text and print the call number in other than the traditional left-hand margin. We submit that one single line at the top of the card might be reserved for the entire call number, followed by 3 lines reserved for headings (tracing). The remainder of the card could then be utilized, with minimum margins, for the text. An alternate possibility is that catalog users be asked to read the call number directly from its position at the bottom of the card. This presupposes that the number be used without change.

B. Cross-reference 3×5 card program. Despite the fact that File 3 and 4 included only new references (i.e., new descriptive headings with their references, or new references to old headings) and that these were not necessarily pertinent to the MARC records, the files promised to be useful in providing current authority file information at an earlier date than such information is available in the published catalogs or in the supplements to the subject heading list. We planned to print multiple sets of 3×5 cards from the tape distributed every other week, since two weeks' references were always to be on a given tape, and distribute these to a number of libraries in the area who expressed an interest in having them. These cards would have been manually interfiled.

The cross-reference 3×5 card program proved unusable, however, since it was unable to continue printing after encountering a continuation card. LC's Information Systems Office reported that the program would not be corrected, and that Files 3 and 4 would be discontinued in favor of providing an improved authority file system at some future date. Accordingly, we wrote a program to cumulate Files 3 and 4 respectively, for possible future use, and abandoned the distribution project.

Provision of authority file information, at the time that the original catalog record is provided, is essential to an effective catalog system, in either card or book form. Perhaps more than any other type of library, a public library wants a work to appear in its catalog, and preferably on the shelves, under the name of the author by which it will be asked for. This is generally the name as it appears on the title page. This is the way the book is advertised, reviewed, and sold. The basic rule for headings for persons in the Anglo-American Cataloging Rules[3] is to "enter a person under the name by which he is commonly identified, whether it is his real name, assumed name, nickname, title of nobility, or other appellation. . . . The form of name of an author . . . is ordinarily determined from the way it appears in his works issued in his language." Since this rule is followed by the Library of Congress only for names not already established under a different form, the result is a basic inconsistency of entry which demands at the very least adequate cross-referencing. This should be provided as an integral part of the MARC system and not require the intellectual intervention of a cataloger and local input of information to the catalog production system.

---

[3] *Anglo-American Cataloging Rules.* Prepared by the American Library Association (et al) North American Text. Chicago, ALA, 1967.

With respect to subject headings, catalog users will seldom be able to come up with a search term that matches the LC subject heading term, and both *see* and *see also* references are essential. The system should be able to generate automatically the appropriate references for any given issue of the book catalog.

VII. Participant Experience in Relation to Computer Facility, Equipment, Programming Staff Problems, etc. Local Programs Written for MARC

When the library was invited to use the IBM 360 computer in the State Department of Institutions, we were offered weekend run time as available, not guaranteed time. This was accepted, mainly because of the better catalog and card printing offered by the expanded print chain. A secondary consideration was a very favorable hourly rate. These considerations seemed to outweigh the disadvantage of the weekend scheduling. However, demand for time on third generation computers by state agencies has grown steadily, and pending installation of such a facility at the Data Processing Service Center, much of the work has had to be channeled to the Department of Institutions. This heavy workload, and disruption of service caused by substitution of a Model 40 for the Model 30, have occasioned run postponements in several instances. The change from 7-level to 9-level tapes upon installation of the Model 40 caused further delays while we arranged for the 7-level tapes to be converted to 9-level tapes in compatibility mode. This conversion is now provided for us by the Argonne National Laboratory.

Irregularity of computer runs resulted in less actual use of the MARC record for current cataloging, as books could not be held up to wait for the MARC record if a proof slip or other source of cataloging data was available. Irregularity also resulted in skewing of the results of tests we attempted to conduct on interval between date of first asking for the MARC record and date of obtaining match. It also affected the results of an effort to analyze relationships between dates of ordering and receipt of the book and dates of availability of proof slip and MARC record. Some processing has been done on the IBM 1401, at the expense of quality of punctuation. (See Figure 5).

Programming staff problems have been nonexistent. The library's programming staff of one has been augmented by the services of an experienced programmer in the Data Processing Service Center, who devoted 750 hours to the author and title catalog, the subject catalog print program, and the label production program. All remaining programming, the coordination of the contract programming, and the arrangements and preparation for all production have been performed by the staff programmer. The time required for programming doubtless reflects lack of experience on the part of the programming staff, especially in handling bibliographic materials, and lack of familiarity with computer technology on the part of the librarian directing the project. Against these lacks, however, must be weighed the interest, enthusiasm, cooperation, and hard work of both programmers.

Programs written locally were:

Procedure	*Description*
MARC–A	Translates characters from LC configuration to local computer. Uses Program ALT 10.
MARC–B	Selects local acquistions from LC tapes. Punches locator cards, caller cards for multiple 3×5 card production for local libraries and caller cards for register print. Creates a tape for update of local holdings and use in production of 3×5 cards. Uses Program MARC I.
MARC–C	This group of programs is divided into author, title, and subject sections. The sections are quite similar in the type of work they do, but handle different information. The first run on the author, title, or subject program uses the newly updated holdings tape. It selects information needed from this tape

to update the author, title, or subject holdings tape. For example, in the title tape, the portions selected would be as follows:

1. A sort field
2. Short title
3. Author information
4. Publisher information
5. LC and Dewey Classification numbers
6. Library holdings information
7. Indicator for Adult, Juvenile, etc.

Step 2 is to sort in author, title, or subject order.

Step 3 is the update of the previous tape with the new tape and any necessary local information, such as *see also* references, etc.

Step 4 prints the catalog from the new tape.

MARC–D	3×5 cards are printed. With switch setting options, it can print main entry, traced entries, and multiple copies of main entry. It uses the original 3×5 program and version A. It also prints the Register in the 3×5 format. The LC tape is used for input.
MARC–E	Creates labels for book preparation materials, using MARC I tape. Uses program SL 012.
MARC–F	Maintains cross-reference files. Uses programs SL 001 and 002.

VIII. Administrative or Managerial Experience with MARC

A. Reaction of library staff.

1. The reaction of the Processing Center staff is that the MARC concept has great potential if we could get MARC catalog records for enough of our titles, as early as we need it (preferably prior to receipt of the book) and have absolutely regular computer time.

2. The reaction of the Timberland Demonstration librarians to the sample catalog is enthusiasm for multiple access points in the author catalog as contrasted with the main entry only, with only occasional references from an added author entry, in the Timberland catalog. No criticisms have been voiced.

3. There have been no adverse comments on the sample catalog from either of the other library systems or from other individuals.

4. There is some objection on the part of State Library staff to the appearance of the catalog cards. Doubtless this is partly due to their being all in upper case, but the chief objection on the part of both the State Library and public library staffs is the frequency of extension cards occasioned by the large type. The custom of shortened catalog cards for shelflist use, as well as for branch use or for secondary cards. seems well established, and two cards are too many.

5. The most serious anti-MARC reaction is that occasioned by the administrative decision to accept the Decimal Classification number in the record without change except for shortening. This decision has not been made by all three systems covered by the book catalog, as well as by the State Library.

A temporary compromise on the question of biography classification allows the library to designate a B for any title it wants to shelve as biography. This B is keypunched as a special designation and precedes the regular subject classification in the book catalog. Where the B is present, that library will ignore the class number and shelve by B + biographee.

Also, we plan to make a study of the problems of adoption of the Decimal Classification expansion of Pacific Northwest history numbers by libraries who have used the expansion published by Charles W. Smith in 1940.[4]

Any objection to the adoption of the Decimal Classification numbers "as is" is an objection to acceptance of centralized classification vs. modification of class numbers to fit an existing collection or an earlier edition of the classification.

This cannot be called an objection to MARC itself; the fact is that the change in policy was occasioned by MARC.

B. Impact of MARC on local automation plans. The proposed library network for Washington State has as a major component a system of book catalogs covering all types of libraries within a given geographic area, and a superimposed network of specialty catalogs. The catalog production system presupposes availability of MARC tapes, with which all original and retrospective input would be made compatible.

---

[4] "An Expansion of the Dewey Decimal Classification for the History of the Pacific Northwest," by Charles W. Smith. *Pacific Northwest Quarterly*, v. 31, no. 2, April 1940.

C. Time frame required to implement a local MARC project. Essentially a year has been required between receipt of the MARC I test tape and completion of the first phase of the system. It is to be noted that by no means has full time been devoted to MARC programming. The contract programmer's time averaged about 88 hours a month for the first six months of the year, and 38 hours a month for the last six months.

D. Other organizational or managerial problems. Having to maintain parallel production of two current catalogs and do most current processing from proof slips or other cataloging sources because MARC coverage is insufficiently broad or up-to-date complicated Processing Center operation and reduced efficiency somewhat. We have tried to take advantage of MARC in every way possible and actually use the product, and have constantly changed procedures as any new capability developed.

IX. Special Studies Done by Participant

A. Performance or Use Studies made of MARC. Figures for overall participant use of the MARC record are given in Table 1. Analyses of availability of needed records and their timeliness were also made periodically.

Table 2, MARC coverage of North Central Regional Library titles for the period December 1, 1966-April 30, 1967, shows coverage of 10% of their titles for the first two months, increasing to a 27% coverage of their titles for the last three months of the period. Overall coverage for five months is 14%, or 19% if only 1966-1967 imprints received are considered.

Table 3 shows the percent of match of Timberland Library Demonstration titles received for the period April 2-May 21, 1967 to vary from 27% to 72% when the titles input and titles matched in a given week are compared.

Table 4 shows that MARC records were found for 39% of the titles cataloged for Timberland during the year 1967.

Table 5 shows the results of a test made of 339 Timberland titles ordered between September 27 and October 13, 1967, to determine availability of the MARC records vs. proof slips with respect to arrival of the books. The results show that we had MARC records for about half as many titles by the date of book receipt as we had proof slips. It also shows that the MARC record was available about two weeks later than the proof slip. If we could schedule our computer runs on the night of receipt of the tape, this delay could be cut almost in half. Furthermore, results show an overall availability of MARC of 72%, given enough time, against the 1967 average of 39%.

B. Cost evaluation studies. A record of MARC participation costs has been maintained as part of our overall cost accounting procedures for the Processing Center.

Systems design, programming, and estimated program testing costs have been separated out as developmental costs. It is impossible to calculate the amount of the remaining production costs that is attributable to learning routines, changing routines, and all of the time it takes to establish a system. Furthermore, other than direct MARC processing is involved in running a cataloging center, such as proof slip searching for participating libraries when MARC records are unavailable.

Because of the variety of products and lack of accounting by program at the computer installation, we have been unable to break down costs by smaller units, such as costs for matching, for printing a 3×5 card, for creating the book catalog tapes, printing a book catalog master, etc. It is therefore equally impossible to get a valid per unit cost, because each operation uses a different number of records.

The so-called production cost of $8999 may reasonably be divided by the number of records input to the system, namely 17,098, since each input record involves clerical, keypunching and computer match time.

This produces a unit cost of $.53. If one considers only the "hits," or those for which a product was obtained (7350), the unit cost goes up to $1.22. Finally, if one considers only the number of unique MARC records used (5712), the unit cost rises to $1.57.

C. Other studies. A study was made of the percentage of titles used in the book catalog for which Decimal Classification numbers were provided in the MARC record. Programming Services, Inc. was interested in whether a tape containing all titles assigned a Decimal Classification number would serve adequately the small and medium sized public library. Out of a total of 12,835 titles on the tape at the time of the study, 510 (4%) were non-fiction titles with no DC number given. Of the 510, the three public library systems selected 10 titles (2%). The latest information from the Library of Congress on Decimal Classification office coverage indicates that the percentage of titles requiring class number assignment should diminish to the vanishing point in the future.

X. Distribution Problems

A. Reaction to air-mailed magnetic tape mode of MARC distribution. This has been normally satisfactory for our purposes. Naturally, distribution of selected records over communication facilities would be preferable when feasible and economical.

B. Recommendations on utility of the MARC data base:

1. How organized and maintained by the user: It is taken for granted that LC cannot continue to cumulate the weekly tapes by card number. Back files may be maintained by card number year. Current input can then be run against known parameters. The author/title tape has not been used by us at all, and is not required.
2. The author/title listing was used only for checking problems, not regularly for checking presence of a title on tape. Its greatest usefulness would be in checking for other card numbers for the same title, e.g., British vs. American editions. A program should be written to print author/title listings from File 1, giving author, short title, publisher, and card number.
3. Comments on the cross-reference file have been made under the section on "Reaction to LC-Supplied MARC Participant Programs."

C. Recommendations for content of the data base. 1st priority: English language monographs with current year card number or publication date, i.e., *Publishers' Weekly* coverage, and omitting English language PL480 cards.
2d priority: Same coverage for retrospective materials working backwards about five years. It will be more useful to use to have retrospective coverage of this class of material than a poor coverage of European materials.
3d priority: Serials. Here new titles would not be the most useful; currently published American periodicals would be our first choice within this category.
4th priority: Current phonorecords.
5th priority: Current educational films.

XI. Conclusions

A. Resultant savings, if any. Cost implications are not really calculable until we can substitute one system for another, rather than utilize a mixture which involves some duplication. It is almost a certainty that we will get more and do more with our money.

B. Resultant improvements in technical processing. Definite work saving in avoiding searching for proof slips, editing, numeric coding, typing and/or reproducing cards, pockets, and spine labels. Allows fuller catalog information for the holding library and more access points in the catalog indexes. Avoids labor of ordering LC cards, waiting for delivery, and typing of headings on card catalog sets.

C. Resultant new products or services.

1. Register section of the book catalog provides full cataloging data without high cost of updating and frequent reprinting.
2. Centralized cataloging service and book catalog production for widely distributed libraries can be undertaken without handling books. The statewide library network could not be considered if every new title required local conversion to machine-readable form. Interlibrary loan will be greatly facilitated as a result of information on location of titles within the state.

D. Other benefits. Standardization of the machine record is essential for communication purposes. Acceptance of standardization of entry, subject headings, and classification is not only more economical, but makes interlibrary cooperation, by means of union catalogs, etc., more attainable. With time and practice, improvements can be made which are not now possible, such as access to all portions of subject headings, automatic analytics, references from computerized authority files, etc.

FIGURE 1.—*Performance or Use Measurement*

Total Participant Use

Number of records on tape, November 1966-December 1967	28082
Total records used by participant, January-December 1967	5712
Percentage of MARC records used to total records on tape	20%

MARC Book Catalog System, January-December 1967

Number of input records punched	14615
Number of times a matching record was found for any of three libraries	5700
Number of unique records matched	4062
Number of 3 x 5 main entry cards printed	6700
Number of titles printed in sample book catalogs	342
Number of entries in each copy of sample book catalog	1444
Percentage of MARC records used to total records on tape	14%

State Library Catalog Card Production, June-December 1967

Number of input records punched (1966 and 1967 titles only)	2483
Number of records matched	1650
Number of 3 x 5 cards printed (sets and extra main entries)	11550
Percentage of MARC records used to total records on tape	7%

FIGURE 2.—*North Central Regional Library MARC Coverage, December 1, 1966—April 30, 1967*

Number of titles received December 1966-January 1967	665
Number of MARC records for December-January titles	69
Percentage of MARC records to December-January titles	10%
Number of titles received February-April 1967	199
Number of MARC records for February-April titles	54
Percentage of MARC records to February-April titles	27%
Total number of titles received December 1966-April 1967	864
Number of titles received with 1966–1967 imprints, December-April	647
Total MARC records, December 1966-April 1967	123
Percentage of MARC records for all December-April titles	14%
Percentage of MARC records for 1966–1967 imprints	19%

FIGURE 3.—*Timberland Library Demonstration MARC Coverage and Timeliness, April—May 1967*

Week	Number of titles received	Number of MARC records matched on subsequent computer run	Percent of match
April 2	274	75	27%
April 9	175	81	46%
April 16-May 1 (combined 3 week run)	582	250	43%
May 7	135	60	44%
May 14	83	60	72%
May 21	243	105	43%

FIGURE 4.—*Timberland Library Demonstration MARC Coverage, Calendar Year 1967*

Number of titles cataloged	6552
Number of MARC records available	2550
Percentage of MARC records available	39%

FIGURE 5.—*Timberland Library Demonstration Test of MARC vs. Proof Slip Availability, September 27, 1967—January 20, 1968*

Number of titles ordered, September 27-October 13, 1967	339
Number of titles received by end of test period (January 20, 1968)	282
Number of proof slips received [1] by end of test period	266
Number of MARC records received [2] by end of test period	217
Number of proof slips received by the date of arrival of book	204
Number of MARC records received by the date of arrival of book	109
Number of titles for which both proof slips and MARC records were received by end of test period	190
Percentage of proof slips received by end of test period	94%
Percentage of MARC records received by end of test period	77%
Percentage of proof slips received by the date of arrival of book	72%
Percentage of MARC records received by the date of arrival of book	38%
Average delay in availability of MARC record beyond availability of proof slip	13.66 days

---

[1] Proof slip receipt date is first date a successful search was made following ordering of the titles, not necessarily the date of receipt of the proof slip in the Processing Center.

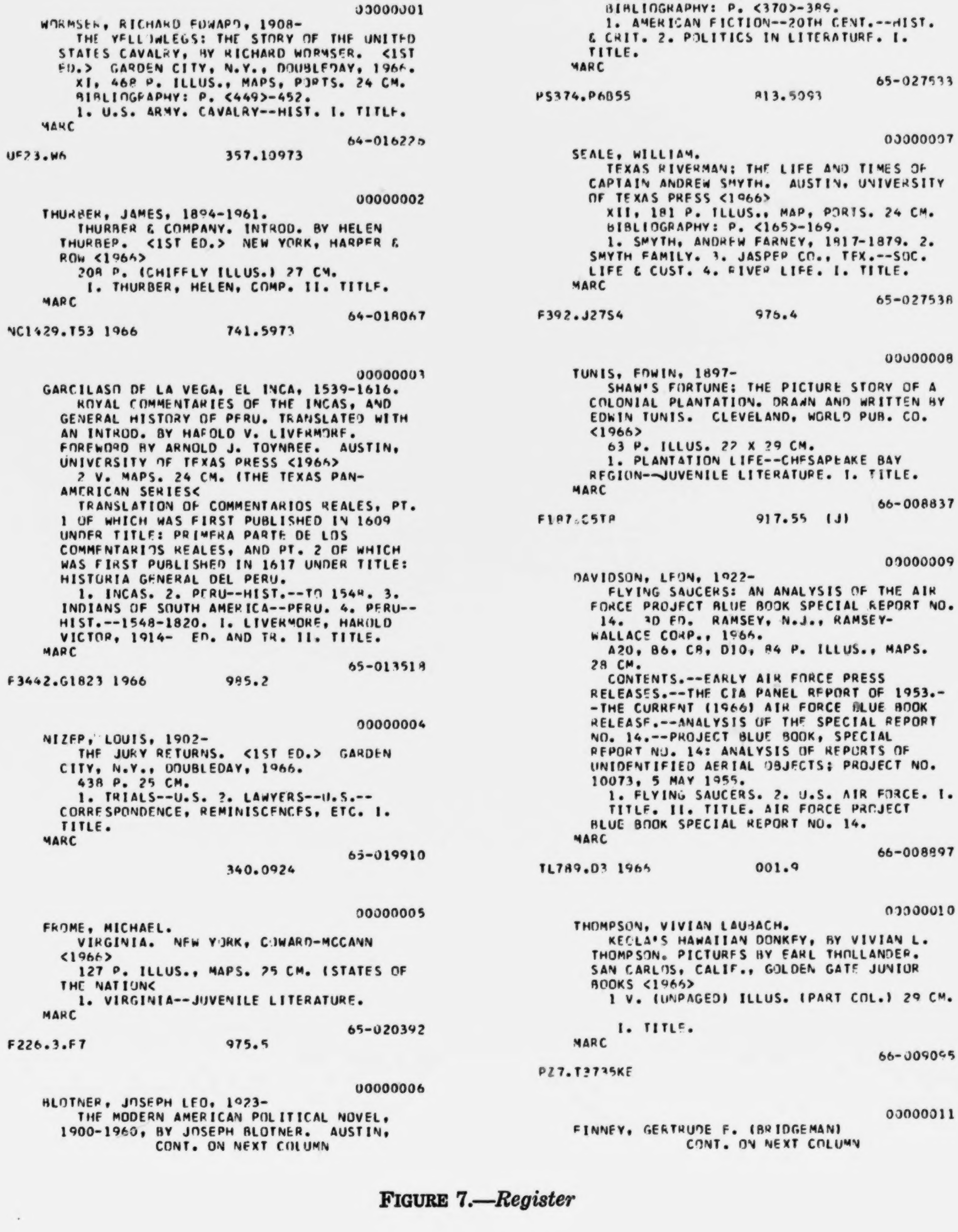
[2] MARC record receipt date is date that 3 x 5 card was available following a successful match to the tape, not the date of the first tape on which it appeared.

FIGURE 6.—*MARC Participation Costs*

	July 1966–June 1967	July-December 1967	Total 18 months
Development Costs:			
System design	3148		3148
Programming:			
Salary	2676	1751	4427
Contract	3426	1448	4874
Program testing	1117	829	1946
Sub-total	10367	4028	14395
Production Costs:			
Keypunching	1175	2854	4029
Clerical operations	665	1125	1790
EAM/EDP time	1953	1659	3612
Supplies	2815		2815
Printing	106		106
Overhead	2391	*	2391
Sub-total	8999	5744	14743
Total	19366	9772	29138
Cost Effectiveness			**Per Unit Cost**
Production Cost	$8999		
Number of records input	17098		.53
Number of records matched	7350		1.22
Number of unique records used	5712		1.57

---

*Overhead cost included in individual activity cost for July-December 1967.



```
                                              00000001
   WORMSER, RICHARD EDWARD, 1908-
      THE YELLOWLEGS: THE STORY OF THE UNITED
     STATES CAVALRY, BY RICHARD WORMSER.  <1ST
     ED.>  GARDEN CITY, N.Y., DOUBLEDAY, 1966.
       XI, 468 P. ILLUS., MAPS, PORTS. 24 CM.
       BIBLIOGRAPHY: P. <449>-452.
       1. U.S. ARMY. CAVALRY--HIST. I. TITLE.
   MARC
                                              64-016225
UF23.W6                   357.10973


                                              00000002
   THURBER, JAMES, 1894-1961.
      THURBER & COMPANY. INTROD. BY HELEN
     THURBER.  <1ST ED.>  NEW YORK, HARPER &
     ROW <1966>
       208 P. (CHIEFLY ILLUS.) 27 CM.
        I. THURBER, HELEN, COMP. II. TITLE.
   MARC
                                              64-018067
NC1429.T53 1966           741.5973


                                              00000003
   GARCILASO DE LA VEGA, EL INCA, 1539-1616.
      ROYAL COMMENTARIES OF THE INCAS, AND
     GENERAL HISTORY OF PERU. TRANSLATED WITH
     AN INTROD. BY HAROLD V. LIVERMORE.
     FOREWORD BY ARNOLD J. TOYNBEE.  AUSTIN,
     UNIVERSITY OF TEXAS PRESS <1966>
       2 V. MAPS. 24 CM. (THE TEXAS PAN-
     AMERICAN SERIES<
       TRANSLATION OF COMMENTARIOS REALES, PT.
     1 OF WHICH WAS FIRST PUBLISHED IN 1609
     UNDER TITLE: PRIMERA PARTE DE LOS
     COMMENTARIOS REALES, AND PT. 2 OF WHICH
     WAS FIRST PUBLISHED IN 1617 UNDER TITLE:
     HISTORIA GENERAL DEL PERU.
       1. INCAS. 2. PERU--HIST.--TO 1548. 3.
     INDIANS OF SOUTH AMERICA--PERU. 4. PERU--
     HIST.--1548-1820. I. LIVERMORE, HAROLD
     VICTOR, 1914-  ED. AND TR. II. TITLE.
   MARC
                                              65-013518
F3442.G1823 1966          985.2


                                              00000004
   NIZER, LOUIS, 1902-
      THE JURY RETURNS.  <1ST ED.>  GARDEN
     CITY, N.Y., DOUBLEDAY, 1966.
       438 P. 25 CM.
       1. TRIALS--U.S. 2. LAWYERS--U.S.--
     CORRESPONDENCE, REMINISCENCES, ETC. I.
     TITLE.
   MARC
                                              65-019910
                          340.0924


                                              00000005
   FROME, MICHAEL.
      VIRGINIA.  NEW YORK, COWARD-MCCANN
     <1966>
       127 P. ILLUS., MAPS. 25 CM. (STATES OF
     THE NATION<
       1. VIRGINIA--JUVENILE LITERATURE.
   MARC
                                              65-020392
F226.3.F7                 975.5


                                              00000006
   BLOTNER, JOSEPH LEO, 1923-
      THE MODERN AMERICAN POLITICAL NOVEL,
     1900-1960, BY JOSEPH BLOTNER.  AUSTIN,
               CONT. ON NEXT COLUMN
```

```
     UNIVERSITY OF TEXAS PRESS <1966>
       X, 424 P. 24 CM.
       BIBLIOGRAPHY: P. <370>-389.
       1. AMERICAN FICTION--20TH CENT.--HIST.
     & CRIT. 2. POLITICS IN LITERATURE. I.
     TITLE.
   MARC
                                              65-027533
PS374.P6B55               813.5093


                                              00000007
   SEALE, WILLIAM.
      TEXAS RIVERMAN: THE LIFE AND TIMES OF
     CAPTAIN ANDREW SMYTH.  AUSTIN, UNIVERSITY
     OF TEXAS PRESS <1966>
       XII, 181 P. ILLUS., MAP, PORTS. 24 CM.
       BIBLIOGRAPHY: P. <165>-169.
       1. SMYTH, ANDREW FARNEY, 1817-1879. 2.
     SMYTH FAMILY. 3. JASPER CO., TEX.--SOC.
     LIFE & CUST. 4. RIVER LIFE. I. TITLE.
   MARC
                                              65-027538
F392.J27S4                976.4


                                              00000008
   TUNIS, EDWIN, 1897-
      SHAW'S FORTUNE: THE PICTURE STORY OF A
     COLONIAL PLANTATION. DRAWN AND WRITTEN BY
     EDWIN TUNIS.  CLEVELAND, WORLD PUB. CO.
     <1966>
       63 P. ILLUS. 22 X 29 CM.
       1. PLANTATION LIFE--CHESAPEAKE BAY
     REGION--JUVENILE LITERATURE. I. TITLE.
   MARC
                                              66-008837
F187.C5T8                 917.55  (J)


                                              00000009
   DAVIDSON, LEON, 1922-
      FLYING SAUCERS: AN ANALYSIS OF THE AIR
     FORCE PROJECT BLUE BOOK SPECIAL REPORT NO.
      14.  3D ED.  RAMSEY, N.J., RAMSEY-
     WALLACE CORP., 1966.
       A20, B6, C8, D10, 84 P. ILLUS., MAPS.
     28 CM.
       CONTENTS.--EARLY AIR FORCE PRESS
     RELEASES.--THE CIA PANEL REPORT OF 1953.-
     -THE CURRENT (1966) AIR FORCE BLUE BOOK
     RELEASE.--ANALYSIS OF THE SPECIAL REPORT
     NO. 14.--PROJECT BLUE BOOK, SPECIAL
     REPORT NO. 14: ANALYSIS OF REPORTS OF
     UNIDENTIFIED AERIAL OBJECTS; PROJECT NO.
     10073, 5 MAY 1955.
       1. FLYING SAUCERS. 2. U.S. AIR FORCE. I.
       TITLE. II. TITLE. AIR FORCE PROJECT
     BLUE BOOK SPECIAL REPORT NO. 14.
   MARC
                                              66-008897
TL789.D3 1966             001.9


                                              00000010
   THOMPSON, VIVIAN LAUBACH.
      KEOLA'S HAWAIIAN DONKEY, BY VIVIAN L.
     THOMPSON. PICTURES BY EARL THOLLANDER.
     SAN CARLOS, CALIF., GOLDEN GATE JUNIOR
     BOOKS <1966>
       1 V. (UNPAGED) ILLUS. (PART COL.) 29 CM.

        I. TITLE.
   MARC
                                              66-009095
PZ7.T3735KE


                                              00000011
   FINNEY, GERTRUDE E. (BRIDGEMAN)
               CONT. ON NEXT COLUMN
```

FIGURE 7.—*Register*

```
ABDUCTION--KILKENNY, IRE. (COUNTY)
                                                 00000319
  WEINER, MARGERY.
    MATTERS OF FELONY.     ATHENEUM, 1967.
  KC    364.1540

ACTING.
                                                 00000123
  MACKENZIE, FRANCES.
    THE AMATEUR ACTOR.          THEATRE ARTS BOOKS
  <1966>
                                          TD    792.028

ACTORS--PORTRAITS.
                                                 00000073
  MCDOWALL, RODDY.
    DOUBLE EXPOSURE.     DELACORTE PRESS <1966>
                       NC    792.0922

ADMINISTRATIVE REMEDIES--U.S.
                                                 00000106
  GELLHORN, WALTER, 1906-
    WHEN AMERICANS COMPLAIN.        HARVARD
  UNIVERSITY PRESS, 1966.
  KC    353

AERONAUTICS--ACCIDENTS.
                                                 00000325
  LEWIS, FLORA.
    ONE OF OUR H-BOMBS IS MISSING.
  MCGRAW-HILL <1967>
  KC    946.082

AERONAUTICS--ACCIDENTS--1966.
                                                 00000186
  MORRIS, CHRISTOPHER, 1938-
    THE DAY THEY LOST THE H-BOMB.
  COWARD-MCCANN <1966>
                                          TD    946.082

AFRICA, NORTH.
                                                 00000082
  STEEL, RONALD, ED.
    NORTH AFRICA.     H. W. WILSON CO., 1967.
  KC   R916.103     NC    916.103

AFRICA, SOUTH--ECON. CONDIT.
                                                 00000029
  HEPPLE, ALEXANDER, 1904-
    SOUTH AFRICA, A POLITICAL AND ECONOMIC
  HISTORY. F. A. PRAEGER <1966>
                       NC    968

AFRICA, SOUTH--HIST.
                                                 00000029
  HEPPLE, ALEXANDER, 1904-
    SOUTH AFRICA, A POLITICAL AND ECONOMIC
  HISTORY. F. A. PRAEGER <1966>
                       NC    968

AFRICA, SOUTH--RACE QUESTION.
                                                 00000029
  HEPPLE, ALEXANDER, 1904-
    SOUTH AFRICA, A POLITICAL AND ECONOMIC
  HISTORY. F. A. PRAEGER <1966>
                       NC    968

AGED--MEDICAL CARE--U.S.
                                                 00000176
  HARRIS, RICHARD O.
    A SACRED TRUST.         NEW AMERICAN LIBRARY
  <1966>
                       NC    610.6273

ALCOHOL AND WOMEN.
                                                 00000087
  KENT, PATRICIA.
    AN AMERICAN WOMAN & ALCOHOL.        HOLT,
  RINEHART AND WINSTON <1967>
  KC    301.4768

ALCOHOLICS.
                                                 00000087
  KENT, PATRICIA.
    AN AMERICAN WOMAN & ALCOHOL.        HOLT,
  RINEHART AND WINSTON <1967>
  KC    301.4768

AMATEUR THEATRICALS.
                                                 00000123
  MACKENZIE, FRANCES.
    THE AMATEUR ACTOR.          THEATRE ARTS BOOKS
  <1966>
                                          TD    792.028

AMERICA--DESCR. & TRAV.--VIEWS.
                                                 00000037
  LEHNER, ERNST, 1895-
    HOW THEY SAW THE NEW WORLD.         TUDOR PUB.
  CO. <1966>
                                          TD J  973.1022

AMERICA--DISC. & EXPLOR.
                                                 00000037
  LEHNER, ERNST, 1895-
    HOW THEY SAW THE NEW WORLD.         TUDOR PUB.
  CO. <1966>
                                          TD J  973.1022

AMERICA--DISC. & EXPLOR.--MAPS.
                                                 00000037
  LEHNER, ERNST, 1895-
    HOW THEY SAW THE NEW WORLD.         TUDOR PUB.
  CO. <1966>
                                          TD J  973.1022

AMERICAN FICTION--20TH CENT.--HIST. & CRIT.
                                                 00000006
  BLOTNER, JOSEPH LEO, 1923-
    THE MODERN AMERICAN POLITICAL NOVEL,
  1900-1960. UNIVERSITY OF TEXAS PRESS <1966>
  KC    813.5093

AMERICAN LITERATURE--HIST. & CRIT.
                                                 00000324
  LEVIN, DAVID, 1924-
    IN DEFENSE OF HISTORICAL LITERATURE.        HILL
  AND WANG <1967>
                                          TD    810.92

AMERICAN LITERATURE--20TH CENT.--ADDRESSES, ESSAYS,
    LECTURES.
                                                 00000215
  COWLEY, MALCOLM, 1898-
    THINK BACK ON US.       SOUTHERN ILLINOIS
  UNIVERSITY PRESS <1967>
                                          TD    810.9005

AMERICAN MEDICAL ASSOCIATION.
                                                 00000176
  HARRIS, RICHARD O.
    A SACRED TRUST.         NEW AMERICAN LIBRARY
  <1966>
                       NC    610.6273
```

FIGURE 8.—*Subjects*

```
                                            00000056
AFFECTIONATELY YOURS, GEORGE WASHINGTON.   NORTON
     <1967>
  WASHINGTON, GEORGE,  PRES. U.S., 1732 -1799.
  KC    973.4109

                                            00000159
AN AGE OF BARNS.  FUNK & WAGNALLS <1966?>
  SLOANE, ERIC.
                    NC    728.9

                                            00000122
ALAMEIN TO ZEM ZEM.  CHILMARK PRESS <C1966>
  DOUGLAS, KEITH CASTELLAIN,  1920 -1944.
                                      TD    940.5481

                                            00000123
THE AMATEUR ACTOR.  THEATRE ARTS BOOKS <1966>
  MACKENZIE, FRANCES.
                                      TD    792.028

                                            00000057
AMERICA'S CONCENTRATION CAMPS.  NORTON <1967>
  BOSWORTH, ALLAN R.
                    NC    940.5472

                                            00000337
AMERICAN AND ENGLISH GENEALOGIES IN THE LIBRARY OF
     CONGRESS.  GENEALOGICAL PUB. CO., 1967.
  U.S.  LIBRARY OF CONGRESS.
                                      TD    016.9291

                                            00000165
THE AMERICAN WAY OF LIFE.  PUTNAM <1967>
  MONTAGU, ASHLEY,  1905 -
  KC    917.3039

                                            00000087
AN AMERICAN WOMAN & ALCOHOL.   HOLT, RINEHART AND
     WINSTON <1967>
  KENT, PATRICIA.
  KC    301.4768

                                            00000032
AMONG THE ANTI-AMERICANS.   HOLT, RINEHART AND
     WINSTON <1967>
  MORGAN, THOMAS BRUCE,  1926 -
  KC    301.2973

                                            00000221
THE ARCHDUKE.  DOUBLEDAY, 1967.
  ARNOLD, MICHAEL P.
                    NC

                                            00000250
ARCHER IN HOLLYWOOD.  KNOPF, 1967.
  MILLAR, KENNETH,  1915 -
                                      TD

                                            00000237
ARCHITECTURAL DRAWING AND LIGHT CONSTRUCTION.
     PRENTICE-HALL <1967>
  MULLER, EDWARD J.
  KC    744.424

                                            00000220
THE ARRANGEMENT.  STEIN AND DAY <1967>
  KAZAN, ELIA.
                    NC

                                            00000311
THE ARROGANCE OF POWER.   RANDOM HOUSE <1967,
     C1966>
  FULBRIGHT, JAMES WILLIAM,  1905 -
                    NC    327.73      TD    327.73

                                            00000222
THE ARTIFICIAL MAN.   PUBLISHED FOR THE CRIME CLUB
     BY DOUBLEDAY, 1967 <C1965>
  DAVIES, LESLIE PURNELL.
                    NC
```

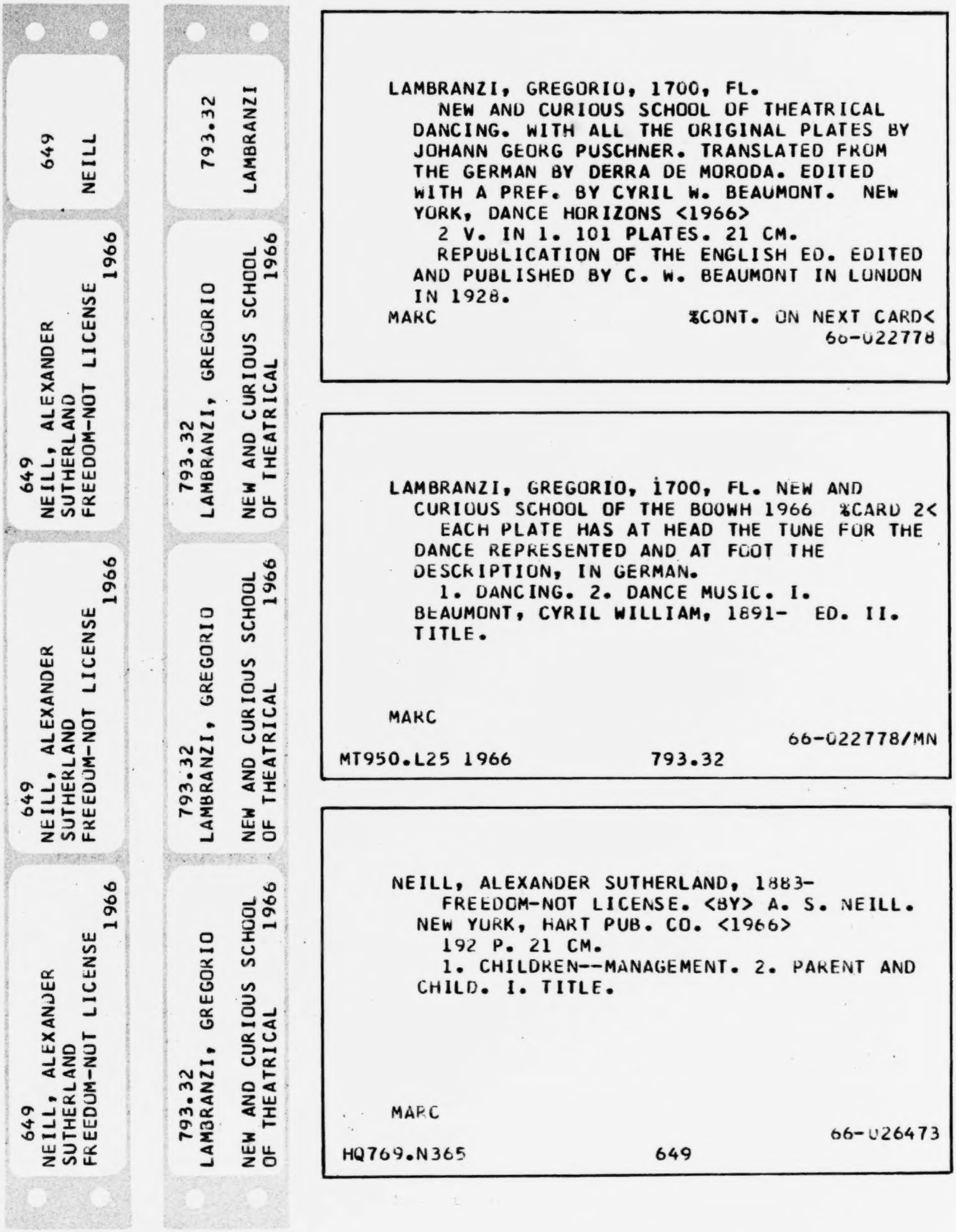
```
                                            00000193
AUSTRALIA AND THE PACIFIC ISLANDS.   H. HAMILTON
     <1966>  5/5/-
  KEAST, ALLEN.
                    NC    500.9099

                                            00000321
THE AUTOBIOGRAPHY OF BERTRAND RUSSELL, 1872-1914.
     LITTLE, BROWN <1967>
  RUSSELL, BERTRAND RUSSELL,  3RD EARL, 1872 -
  KC    828.9120

                                            00000138
BARRON'S GUIDE TO THE TWO-YEAR COLLEGES.   BARRON'S
     EDUCATIONAL SERIES <1966>
  ESKOW, SEYMOUR.
                                      TD    378.1543

                                            00000173
BARRON'S PROFILES OF AMERICAN COLLEGES.   BARRON'S
     EDUCATIONAL SERIES, INC. <1966>
  FINE, BENJAMIN,  1905 -
  KC   R378.0257

                                            00000164
BASEBALL.  PRENTICE-HALL <1967>
  COOMBS, JOHN WESLEY.
  KC    796.357

                                            00000113
BEWARE OF THE BOUQUET.   PUBLISHED FOR THE CRIME
     CLUB BY DOUBLEDAY, 1966.
  AIKEN, JOAN,  1924 -
                    NC

                                            00000235
THE BIG BLUE LINE.  COWARD-MCCANN <1967>
  CRAY, ED.
                    NC    364.1209

                                            00000205
THE BITTER HERITAGE.   HOUGHTON MIFFLIN, 1967
     <C1966>
  SCHLESINGER, ARTHUR MEIER,  1917 -
                    NC    327.7305

                                            00000244
BLACK IS BEST.  PUTNAM <1967>
  OLSEN, JACK.
                    NC    796.8309

                                            00000223
THE BLUE DANUBE COOKBOOK.  DOUBLEDAY, 1967.
  DONOVAN, MARIA KOZSLIK.
  KC    641.594

                                            00000058
BLUEBERRY CULTURE.   RUTGERS UNIVERSITY PRESS
     <1966>
  ECK, PAUL,  1931 -  ED.
                                      TD    634.737

                                            00000171
BOTTLE RUSH U.S.A..   OLD TIME BOTTLE PUB. CO.
     <1966>
  BLUMENSTEIN, LYNN.
  KC    748.8

                                            00000175
BRANDS, TRADEMARKS, AND GOOD WILL.   MCGRAW-HILL
     <1967>
  MARQUETTE, ARTHUR F.
                    NC    338.7633

                                            00000211
BUCCANEERS & PIRATES OF OUR COASTS.   MACMILLAN
     <1967>
  STOCKTON, FRANK RICHARD,  1834 -1902.
                                      TD J  910.453
```

FIGURE 9.—*Titles*

```
                                        00000113
AIKEN, JOAN,  1924 -
    RFWARE OF THE BOUQUET.  PUBLISHED FOR THE
    CRIME CLUB BY DOURLFDAY,  1966.
                  NC

                                        00000192
ALRRAND, MARTHA.
    A DOOR FELL SHUT.  HODDER & STOUGHTON
    <1966>   -/16/-
                  NC

                                        00000037
ALEXANDFR, GERARD L.,  ED.
  LEHNER, ERNST,  1895 -
    HOW THEY SAW THE NEW WORLD.  TUDOR PUR.
    CO.  <1966>
                                   TD J  973.1022

                                        00000182
AMERICAN ASSEMBLY.
    A WORLD OF NUCLEAR POWERS?  PRENTICE-HALL
    <1966>
                                   TD    327.1

                                        00000136
ANDERSON, SHERWOOD,  1876 -1941.
    WINESBURG, OHIO.   VIKING PRESS  <1966>
  KC

                                        00000167
ANNER, GFORGE F.
    ELEMENTARY NONLINEAR ELECTRONICS CIRCUITS.
    PRENTICE-HALL,  <C1967>
  KC    621.3815

                                        00000243
ARMY TIMES,  WASHINGTON, D.C.
    WARRIOR; THE STORY OF GENERAL GEORGE S. PATTON.
    PUTNAM  <1967>
  KC    355.3310

                                        00000221
ARNOLD, MICHAEL P.
    THE ARCHDUKE.   DOUBLEDAY,  1967.
                  NC

                                        00000218
ASCH, SIDNEY H.
    POLICE AUTHORITY AND THE RIGHTS OF THE
    INDIVIDUAL.   ARCO  <1967>
  KC    343.0973

                                        00000291
AUCHINCLOSS, LOUIS.
    TALES OF MANHATTAN.  HOUGHTON MIFFLIN,
    1967.
  KC

                                        00000183
AVIS, WARREN E.,  JOINT AUTHOR.
  BLAKE, ROBERT ROGERS,  1918 -
    CORPORATE DARWINISM.  GULF PUB. CO.
    <1966>
  KC    658

                                        00000197
BALCHIN, NIGEL,  1908 -
    IN THE ABSENCE OF MRS. PETERSEN.
    COLLINS,  1966.   -/21/-
                  NC

                                        00000232
BARLOW, JAMES,  1921 -
    ONE MAN IN THE WORLD.  SIMON AND SCHUSTER
    <C1966>
                  NC
```

```
                                        00000131
BARROW, RHODA,  JOINT AUTHOR.
  COHEN, BEN.
    CALLING A SPADE A SPADE.  A. S. BARNES
    <1966, C1965>
                                   TD    795.4152

                                        00000336
BARTLETT, CHARLES,  JOINT AUTHOR.
  WEINTAL, EDWARD,  1901 -
    FACING THE BRINK.   SCRIBNER  <1967>
  KC    327.73

                                        00000150
BASSITY, MATTHEW A. R.
    THE MAGIC WORLD OF ROSES.  HEARTHSIDE
    PRFSS  <1966>
                                   TD    635.9333

                                        00000279
BATES, DAISY,  1870 -1951.
    THE PASSING OF THE ABORIGINES.  PRAEGER
    <1967>
                                   TD    919.4034

                                        00000140
BEARSS, EDWIN C.
    HARDLUCK IRONCLAD.  LOUISIANA STATE
    UNIVERSITY PRESS,  1966.
                                   TD    359.325

                                        00000096
BEAUMONT, CYRIL WILLIAM,  1891 -  ED.
  LAMBRANZI, GREGORIO,  1700 , FL.
    NEW AND CURIOUS SCHOOL OF THEATRICAL DANCING.
    DANCE HORIZONS  <1966>
                                   TD    793.32

                                        00000299
BELL, GERALD D.,  ED.
    ORGANIZATIONS AND HUMAN BEHAVIOR.
    PRENTICE-HALL  <1967>
  KC    301.18

                                        00000112
BERNANOS, GEORGES,  1888 -1948.
    MOUCHETTE.  HOLT, RINEHART AND WINSTON
    <1966>
                  NC

                                        00000108
BERSON, ABNER,  1946 -  JOINT AUTHOR.
  BERSON, FRED.
    SECRETS OF A PROFESSIONAL COIN DEALER.
    POCKFT BOOKS  <1966>
                                   TD    737.4973

                                        00000108
BERSON, FRED.
    SECRETS OF A PROFESSIONAL COIN DEALER.
    POCKET BOOKS  <1966>
                                   TD    737.4973

                                        00000241
BETTFLHEIM, BRUNO.
    THE EMPTY FORTRESS.   FREE PRESS  <1967>
  KC    618.9289

                                        00000013
BETTER HOMES AND GARDENS.
    TREASURY OF CHRISTMAS IDEAS.  MEREDITH
    PRESS  <1966>
                  NC    394.268

                                        00000310
BIBLE.  O.T. PSALMS. ENGLISH. SELECTIONS. 1967.
     AUTHORIZED.
    SONGS OF JOY FROM THE BOOK OF PSALMS.
    WORLD PUB. CO.  <1967>
                                   TD    223.2
```

FIGURE 10.—*Authors*



```
649
NEILL
```

```
649
NEILL, ALEXANDER
SUTHERLAND
FREEDOM-NOT LICENSE
                1966
```

```
649
NEILL, ALEXANDER
SUTHERLAND
FREEDOM-NOT LICENSE
                1966
```

```
649
NEILL, ALEXANDER
SUTHERLAND
FREEDOM-NOT LICENSE
                1966
```

```
793.32
LAMBRANZI
```

```
793.32
LAMBRANZI, GREGORIO

NEW AND CURIOUS SCHOOL
OF THEATRICAL     1966
```

```
793.32
LAMBRANZI, GREGORIO

NEW AND CURIOUS SCHOOL
OF THEATRICAL     1966
```

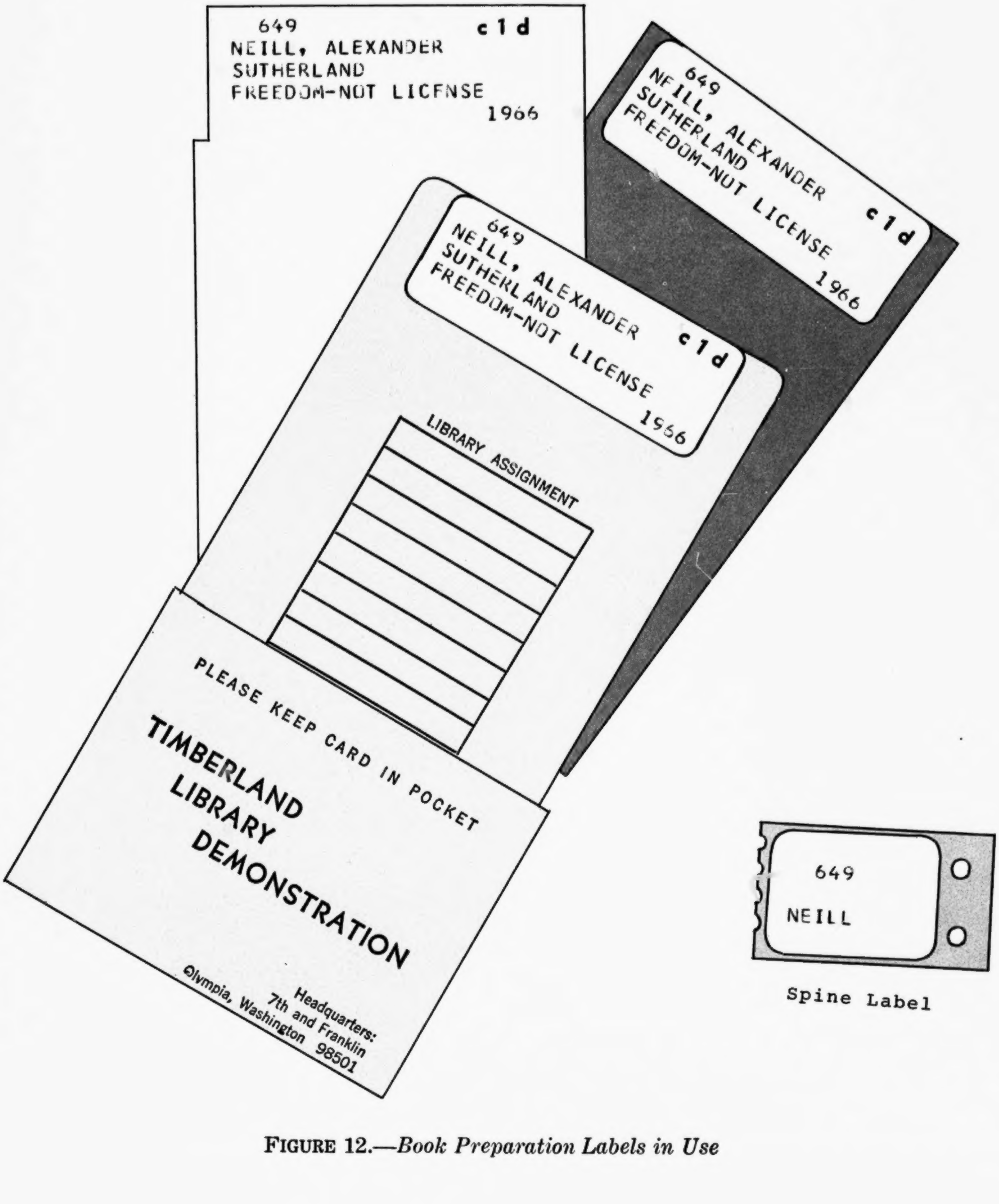
```
793.32
LAMBRANZI, GREGORIO

NEW AND CURIOUS SCHOOL
OF THEATRICAL     1966
```

```
LAMBRANZI, GREGORIO, 1700, FL.
     NEW AND CURIOUS SCHOOL OF THEATRICAL
  DANCING. WITH ALL THE ORIGINAL PLATES BY
  JOHANN GEORG PUSCHNER. TRANSLATED FROM
  THE GERMAN BY DERRA DE MORODA. EDITED
  WITH A PREF. BY CYRIL W. BEAUMONT.  NEW
  YORK, DANCE HORIZONS <1966>
    2 V. IN 1. 101 PLATES. 21 CM.
    REPUBLICATION OF THE ENGLISH ED. EDITED
  AND PUBLISHED BY C. W. BEAUMONT IN LONDON
  IN 1928.
MARC                      %CONT. ON NEXT CARD<
                                     66-022778
```

```
LAMBRANZI, GREGORIO, 1700, FL. NEW AND
  CURIOUS SCHOOL OF THE BOOWH 1966  %CARD 2<
    EACH PLATE HAS AT HEAD THE TUNE FOR THE
  DANCE REPRESENTED AND AT FOOT THE
  DESCRIPTION, IN GERMAN.
    1. DANCING. 2. DANCE MUSIC. I.
  BEAUMONT, CYRIL WILLIAM, 1891-  ED. II.
  TITLE.



MARC
                                  66-022778/MN
MT950.L25 1966          793.32
```

```
NEILL, ALEXANDER SUTHERLAND, 1883-
     FREEDOM-NOT LICENSE. <BY> A. S. NEILL.
  NEW YORK, HART PUB. CO. <1966>
    192 P. 21 CM.
    1. CHILDREN--MANAGEMENT. 2. PARENT AND
  CHILD. I. TITLE.




  MARC
                                     66-026473
HQ769.N365              649
```

FIGURE 11.—*3 × 5 Main Entry Cards and Book Preparation Labels*


649
c l d
NEILL, ALEXANDER
SUTHERLAND
FREEDOM-NOT LICENSE
1966
LIBRARY ASSIGNMENT
PLEASE KEEP CARD IN POCKET
TIMBERLAND
LIBRARY
DEMONSTRATION
Headquarters:
7th and Franklin
Olympia, Washington 98501
649
NEILL

Spine Label

FIGURE 12.—*Book Preparation Labels in Use*

YALE UNIVERSITY

David L. Weisbrod
Head, Development Department, Research and Development

Introduction

This report is concerned with the experience of Yale University Library as a participant in the MARC (Machine-Readable Cataloging) Pilot Project of the Library of Congress. All of the systems and programming work involved in MARC utilization at Yale was performed by the Development Department of the Office of Research and Development, Yale University Library.

Development Department personnel involved in this participation were: Dr. George Crawford, who designed and implemented MATE (*MA*RC *T*ranslate and *E*dit) and has had a hand in almost all of the MARC-related programs; Mrs. Susan W. McNair, who implemented the first version of the call number search program; Miss Sandra E. Stone and Mr. Curtis E. Higgins, who are responsible for the Yale Bibliographic System, with which MARC interfaces through MATE; and Mr. David L. Weisbrod, Head, Development Department. Also involved was Mr. Frederick G. Kilgour, formerly Associate Librarian for Research and Development, Yale University Library.

Other library personnel involved as users of MARC-related products were: Mr. John A. Harrison, Librarian, Kline Science Center; Mr. M. L. Montee, Head, Cataloging Department, Yale Medical Library; Mr. Stanley D. Truelson, Jr., Librarian, Yale Medical Library; and Mr. Bret Slava Zadera, Science Bibliographer, Yale University Library, and Assistant Librarian, Kline Science Center.

A. Description of User Library and Computer Configuration Used

A.1. The Library. Yale University Library is in fact a library system, consisting of the central University Library and more than 60 school, college, and department libraries. The central University Library collection numbers more than three million volumes; the holdings of the entire system exceed five million volumes.

The particular libraries concerned in one way or another with the MARC Pilot Project are Kline Science Center Library (approximately 84,000 volumes) and Yale Medical Library (approximately 382,000 volumes).

A.2. Computer Configuration. Access was available to two computers, an IBM 1401 and an IBM DCS. Both of these machines are at the Yale Computer Center. The library pays the prevailing rate to purchase time on these machines.

The 1401 has 12,000 locations of core memory, two model 729–V tape drives, and the following special features: advanced programming, column binary, direct data channel and associated analog-to-digital converter (not used in conjunction with MARC), high-low-equal compare, multiply-divide, print storage, sense switches, space suppression, and modified character set. The last feature is used in conjunction with an upper/lower case print train for the 1401 printer. Current rate for the 1401 is $20 per hour.

The DCS is a Direct Coupled System comprising an IBM 7040 and an IBM 7094. User programs operate on the 7094 very much as if it were a stand-alone machine. The 7040 is used to stack jobs for the 7094 and to handle the input/output operations for the 7094. The peripheral equipment to which user programs have access, either implicit or explicit, are: one 1402 card read-punch; two 1403 printers (1100 lpm); eight tape drives (729–V and 729–VI); a 1301–2 disk storage device, which is used for both systems, residence and temporary user storage; 740 cathode ray tube recorder and 790 cathode ray tube display (not used in conjunction with MARC). Current rates for the DCS are $275/hour days and evenings, $175/hour nights, i.e., after 1:00 A.M.

B. Original Plans for Use of MARC

The original plan for MARC utilization at Yale consisted of running a new titles alerting service for the Science Bibliographer (book selection librarian).

Because the MARC Pilot Project as originally conceived was to last only six months, it did not seem likely that sufficient time would be available to build and operate an effective interface between MARC and YBS. YBS is the Yale Bibliographic System, a computer-assisted catalogue system that performs the functions of a master file maintenance on magnetic tape, 3×5 catalogue card production, and periodic accessions list production.

C. Actual Implementation Accomplished

Functions actually implemented were (1) a New Titles Alerting Service, (2) an interface MARC and YBS called MATE, and (3) a number of miscellaneous ancillary programs concerned mainly with MARC file maintenance. Each of these functions is treated below in a separate subsection.

C.1. New Titles Alerting Service. This service is provided by searching each week's MARC tape for new titles in specific areas, based on the class number component of the Library of Congress call number (tag 90). Whenever a "hit" is made, the contents of the record are printed out using a slightly modified MARC Bibliographic Listing program. This entire operation is performed on the 1401. The library's Science Bibliographer has been receiving this service since November 1966 (the start of the MARC Pilot Project). The Medical Librarian has been receiving this service, using a different set of selection criteria, since November 1967.

Three versions of the class number search program presently exist, differing only in degree of flexibility and specificity with which the desired classifications can be specified.

C.1.a. Constraints. Each version of the search program has its own particular limit as to the maximum number of classes or subclasses that can be requested.

C.1.b. Operating Problems Encountered. The first version of the search program punched out call cards, containing LC card numbers, for input to the MARC Bibliographic Listing program. This required two passes through the tape, one for searching and one for printing. The search logic was later built into the Bibliographic Listing program, so that only one pass through the tape was necessary.

A separate pass through the tape is necessary for each search request. With only two standing requests, this is not very time-consuming. If the number of customers for this service increases, it will be necessary to implement some sort of combined search logic, which will undoubtedly force the removal of this application from the 1401 to the DCS, at least for all phases other than the final printing.

C.1.c. Results Achieved. The Science Bibliographer has been generally quite pleased with the service, except for two drawbacks which have been identified. The first is that on the average a title appears in the selected MARC listing about a week after the corresponding LC proof sheet is received. The second is that the limitation of the MARC Pilot Project to English language materials leaves uncovered much important foreign language material. Since the LC proof sheets are not organized by language, there is no way of isolating the materials that are not to be covered by MARC. Thus, the MARC-based New Titles Alerting Service is, for the Science Bibliographer, only a close second to the proof sheet service. The time delay is regarded as tolerable and of secondary importance compared with the language problem.

The New Titles Alerting Service better fits the needs of the Medical Librarian. His library is concerned primarily with English language materials, which is where MARC coverage lies. (The Historical Library section, which deals with older works, many of which are in foreign languages, is totally outside the scope of MARC.) In terms of timeliness, the Medical Librarian feels that the MARC-based service competes favorably with the *Current Catalog* of NLM and with *American Book Publishing Record,* which he has routinely used as accessions checklists. It was in response to the needs of the Medical Librarian that the flexibility of the search program's selection logic was increased, to allow the selection of small subclasses or intervals of sub-classes in the LC classification number. The present degree of selectivity is considered adequate.

C.2. MATE (*MA*RC *T*ranslate and *E*dit). The MATE program allows the MARC tapes to be used as a source of direct input to the Yale Bibliographic System. MATE performs the two functions of translating explicitly the content of selected records from MARC format into YBS input format and of editing this translated material according to specific instructions provided by a cataloguer. The translated, edited material then enters YBS exactly as though it were the product of original cataloging.

MATE operates in the following general manner. The inputs to MATE are the MARC tapes and a card deck punched from material supplied by the cataloguer. This card deck is comprised of a number of smaller decklets, each representing a transaction, i.e., each identifying a MARC record to be translated and specifying the editing operations to be performed. The output from MATE is a magnetic tape containing data in a format identical to YBS original cataloguing input.

The cataloguer's source of information in preparing inputs for MATE is a MARC Bibliographic Listing for the work being catalogued. This is obtained either (1) via the New Titles Alerting Service (see Section C.1., above), or (2) in response to a specific request for the listing, keyed on LC card number, i.e., the MARC record identification. In general, a specific request for a MARC Bibliographic Listing is entered when the work being catalogued was not ordered in response to the New Titles Alerting Service but has been located in either the MARC LC card number listing or the MARC abbreviated author/title listing. Through MATE the cataloguer can effect the changing or deletion of any specific category of MARC data, i.e., of any variable field, and the addition of new data to the translated record. The original record on the MARC tape always remains intact.

At the time of this writing YBS is being used in only one library in the Yale University Library system, Yale Medical Library. Thus the medical library is the only place where MATE can be employed. Procedures for using MATE in conjunction with the New Titles Alerting Service were introduced at Yale Medical Library in November 1967.

C.2.a. Constraints. Cost estimates for programmed conversion of the MARC fixed field codes into the YBS leader codes indicated that it would be less costly, at least initially, for the cataloguer to enter the codes "from scratch." The MARC Bibliographic Listing does, of course, display the original MARC codes, thus providing a helpful suggestion for the YBS cataloguer.

C.2.b. Operating Problems Encountered. In the case of the new titles alerting service the only data of interest are the new records on each week's MARC tape. Once the new titles listings have been made the data can, in effect, be thrown away; there is no need for a cumulative MARC master file. The reason for maintaining this file is essentially the same as the reason a library may decide to maintain an LC card depository catalogue: Any book may be acquired at any time, and the only way to have specific LC cataloguing available when it is needed is to keep it all. The problem of maintaining and using a cumulative MARC master file has not been an overwhelming one, but the file has not yet grown to significant proportions. The real problem lies in the future, when MARC II coverage increases and the distribution tapes become non-cumulative. Programs concerned with the maintenance and use of the cumulative MARC file are discussed in Section C.3.

C.2.c. Results Achieved. The general reaction at Yale Medical Library to MATE has been an enthusiastic one. Although the draft version of the cataloguer's instructions appeared formidable at first glance, experience soon revealed that the MATE input notation is easy to use, as we hoped it would be. It is still too early to determine whether MATE is more accurate or economical than the alternative manual transcription of the same data from LC cards or printed catalogue onto YBS worksheets.

C.3. Miscellaneous Ancillary Programs. In order to facilitate both program operation and cataloguer effort, a number of programs have been produced to perform such processes as (1) merge the several files no. 1 from the full MARC tape reels into a single file; (2) edit file no. 2 into a fixed record length, upper case only format, ready for sorting and/or merging with a previously processed file no. 2; (3) extract data from file no. 1 to construct a file no. 2 in the same format as (2).

It was desired to provide the cataloguers at Yale Medical Library with a merged author/title index to the full MARC tape reels. When an attempt was made to merge and print the several files no. 2 on the three completed MARC tapes, it was discovered that: (1) the records were not actually in any sort of reasonable alphabetic sequence (some records had been sorted on diacriticals, for instance), and (2) the records were not all the same length. The files could not be merged "as is." Hence the need for the file no. 2 edit program.

The total absence of a cumulated file no. 2 on MARC reel no. 4 has necessitated the extraction program. This facility allows the production from time to time of a cumulated author/title index to the "live" reel of MARC data, so that the cataloguer does not have to search a large number of noncumulative weekly indexes.

D. Reaction to MARC I Format

Since the principal purpose for any commentary on MARC is to affect the design of MARC II, since extensive conferences concerning this matter have already taken place between LC/ISO and Yale personnel, and since the MARC II format is already rather firmly established, the size of the present section has been kept to a minimum.

D.1. Uses Made of Fixed Fields. The only fixed fields of which use has been made at Yale are field no. 3, Library of Congress catalogue card number, and field no. 25, length of record. No particular use was devised for the other 23 fixed fields. With respect to those fixed fields for which there exists a counterpart in the YBS master record format, it was decided that it was more reasonable, at least for a short-lived experiment, not to implement extensive table-driven code conversion routines.

D.2. Special Uses Made of Variable Fields. None.

D.3. Local-Use Data Elements Employed. None. Rather than add local-use data elements to existing MARC records, it was decided to translate MARC records to YBS format (see Section D.4).

D.4. Modifications Made Locally to Format. MARC records of permanent interest are converted into YBS format using MATE.

E. Reaction to LC-Supplied MARC Participant Programs

The only LC-supplied program of which any serious use was made was the Bibliographic Listing Program. There were a few reproducer (punch) errors in the deck that was delivered to Yale; it was not very difficult to find and correct them.

More annoying, however, was the fact that no use was made in the program of either index registers or sense switches. The first modification made to the program was to put the forms alignment logic under sense switch control. The fact that index registers were not used meant that a primitive address replacement technique was used throughout, which made the program considerably harder to modify, as well as slower to operate. As is already well known at LC/ISO, these two restrictions were totally unnecessary, as the index and sense switch hardware features were available to all MARC participants with IBM 1400 series equipment.

At the time that the Bibliographic Listing Program was being written, the contractor knew that different MARC participants had different print chains. Yet the printing routines were so coded as not to facilitate local character conversion. Generalized, table-driven print logic, such as was employed in Yale's CHY6 program a full year before the inception of the MARC Pilot Project, and of whose existence both the contractor and ISO were aware, would have been vastly superior.

A major "bug" that had to be corrected at Yale was that the program did not correctly compare the LC card number on the call card with the LC card number in the tape record. No allowance had been made for the fact that the prefixes were in uppercase notation in the tape record but were in caseless, i.e., lowercase, notation on the call card.

In spite of these complaints, the MARC Bibliographic Listing Program, as modified locally (see Section C.1.b), has worked reliably at Yale for approximately 15 months.

F. Experience in Relation to Computer Facility, Equipment, Programming Staff Problems, Etc.

The Yale University Library is one of the largest users of the Yale Computer Center. Since MARC utilization is one of the library's smaller application areas, there has been no difficulty in obtaining machine time for MARC work.

With respect to equipment, the only major problem is one that affects the library's YBS work as well as the MARC work. When the library's 120-character chain is not in use on the 1401's 1403 printer, the computer center's 48-character is normally installed. It has proved very difficult—almost impossible—to adjust the 1403 printer so that both the 48- and 120-character chains give good impressions within the tolerances of adjustment of the print density and forms thickness controls which are normally used to control the impression.

The library has its own programming staff in the Development Department. Thus there has been no problem of obtaining dedicated programming assistance from the computer center, with which some of the other participant libraries have had to contend. The Development Department has had one programmer devoted almost full-time to MARC since late fall of 1966.

G. Administrative/Managerial Experience with MARC

G.1. Staffing Problems. We are aware of no staffing problems related to the MARC Pilot Project. The reaction to those members of the library staff who have been involved with the project has been enthusiastic. (see Section C).

G.2. Impact of MARC on Local Automation Plans. All of the work thus far can be described as a feasibility demonstration. This is true for both the new titles alerting service and the MATE function. We hope to be able before the end of the project to gather sufficient data to allow us to perform at least a rough cost analysis for the MARC-related services to determine if in any way they are "paying off."

The cost of MARC-related development and services to date has been absorbed into the general budget of the Development Department. MARC II has not been explicitly allowed for in future budget projections beyond a small investigative allotment because, in part, future costs are very far from determined.

G.3. Time Required to Implement a Local MARC-Related Project. The original selection program for the new titles alerting service required approximately one man-week to write and debug. This effort was spread over two or three calendar weeks.

The MATE facility required approximately seven months to implement (roughly December 1966 through June 1967). But owing to a myriad of conflicting demands on our time and energies, it was not adopted at Yale Medical Library until November 1967.

Time information for the other programs is not available.

H. Special Studies Done by Participant

None.

I. Distribution Problems

I.1. Reactions to Air Mailed Magnetic Tape. Contrary to expectations, this mode of MARC tape distribution was entirely satisfactory. We received not one tape in unreadable condition. And, as far as can be told, delivery was rapid and sufficiently reliable.

I.2. Recommendations on Utility of the MARC Data Base

I.2.a. Organization and Maintenance by the User. It has been found most convenient for our particular needs to use the current ("live") reel intact as delivered, and to combine files from the "frozen" reels. Presently the combined files no. 1 from reels 1, 2, and 3 fit onto one reel, at 800 bpi, blocked into records of 460 six-character words. In similar manner, the combined files no. 2 fit onto another reel. This data compression is important when it is observed that with just one back-up copy, each additional reel of data requires two reels of tape. (With two levels of back-up, the cost factor rises to 3; etc.)

I.2.b. Cumulative versus Non-Cumulative Distribution. There is no single best answer to this question. If the user is in a position to use the MARC tape exactly as delivered as a permanent reference file, then it is probably less costly to the nation as a whole for the file to be cumulated centrally, i.e., for distribution to be cumulative over one reel of tape, as is presently done, rather than non-cumulative. But if for any reason the tape is not to be used exactly as delivered as a permanent reference file, perhaps because new records are to be transferred to disc, because no permanent reference file is to be maintained, or because only selected records are to be retained (say, only non-juvenile materials in French, German, and English), then the benefit of central cumulation is less clear, and it is probably less costly overall to distribute a non-cumulated tape.

The problem is that no matter which way the decision is made the exceptional user gets "punished" by having to incur higher processing costs: The user who only prints a list of new titles from a cumulative tape has to pay the price of skipping over all the old records. The user who cumulates on tape a complete file from a non-cumulative tape has to duplicate the entire current-reel-to-date just for the sake of adding a few records.

A possible solution to this dilemma is to offer two services, perhaps at different prices, one for a non-cumulative tape, the second for a cumulative tape.

I.2.c. Value of the Author/Title File. For a user producing just a new titles listing from the MARC tape the value of the author/title file is nil. But for a user who is maintaining a complete permanent reference file of MARC data, for use in cataloguing, the author/title listing provides a useful and valuable access route to specific records. It is obviously easier and less expensive for the author/title file to be extracted and sorted once centrally for mass distribution than for it to be extracted and sorted many times over at the user libraries.

Two conditions must be met for the author/title file to be truly useful. The first is that the author/title file must be so constructed as to permit the merging of two such files into a single combined file with a simple merge program. (This was not the case with MARC Pilot Project author/title file. See Section C.3.) The second condition is that the user must be retaining the entire MARC main data file for permanent reference. If the user is retaining only a subset, it becomes almost impossible to simultaneously select for retention the corresponding author/title records since they contain so little data. (If this is the case the user has little choice but to extract his own author/title file from his own 'tailored" MARC main data file.)

I.2.d. Value of a Cross-Reference File. The value of a cross-reference file is very much dependent on local programming and file maintenance. Such a file would be extremely valuable to a bookform catalogue production project. Its value to a card catalogue maintenance project is less obvious, as extensive records must first be generated concerning what cross-reference entries already exist in which catalogues, so that program logic may be implemented to automatically produce a new cross-reference card when needed for a specific catalogue.

J. Conclusions

J.1. Resultant Savings. As mentioned in Section G.2, the work done at Yale to date has been essentially a demonstration of feasibility. When cost data have been assembled they will be submitted to the Library of Congress as a supplement to the present report.

J.2. Resultant Improvements in Library Technical Processes Performances. As was suggested in Section C.1., the New Titles Alerting Service probably has not made much difference to either the quality or speed of book selection and ordering for the Kline Science Center Library, whereas it probably has for Yale Medical Library.

In the cataloguing area it is already obvious that both cataloguer's and keypuncher's time is saved by notating only the editing changes necessary to adapt a MARC record to local needs, as compared with both original cataloguing and manual transcription (say, from the printed NUC) of non-machine-readable cataloguing information. The broader question, however, is: Do the necessary supporting operations of MARC search and MATE cost so much as to dissipate the dollar value of this saving? As mentioned above, this is yet to be ascertained. And even if it were to develop that the overall cost of using MARC is greater, could that increase be justified either on the grounds that it results in faster overall cataloguing service or on the grounds that it allows the library to get more production out of its rarest of resources, the trained cataloguer? This is a value/utility judgment that can be made only by the librarian in charge of the allocation of resources.

J.3. Resultant New Products or Services Not Previously Feasible. The New Titles Alerting Service offers a degree of selectivity not attainable through the broad groupings into which the LC proof sheet service organizes its output. This service for the bibliographer can be described fairly as not previously available. But whereas this service may be new to libraries as an internal service, the general concept is far from novel, for this is nothing but the familiar SDI (selective dissemination of information) service in a new context.

Although the New Titles Alerting Service may indirectly improve the library's service to the patron, it is not itself a specific product or service of which the patron may avail himself.

J.4. Other Benefits. None.

J.5. Disadvantages. The main "disadvantage" of MARC is that it costs money to maintain locally the tapes and the listings that serve as access routes to the data on tape. But, to set this "disadvantage" in perspective, the cost of this MARC file maintenance must be compared against the cost of maintaining a comparable depository catalogue of LC cards. Valid data are not available on the basis of which this comparison can be made. One might argue that a better comparison would be between the local cost of MARC file maintenance and the cost of a subscription to the printed NUC service (monthly, quarterly, etc.).

The essential point here is that at Yale unexpected costs were incurred in maintaining a usable MARC file. But the truism, "One never gets something for nothing" is especially relevant here. The policy of minimum cost is to do nothing. The paradox, however, is that to do nothing is not a realistic or reasonable policy, at least as regards libraries and "automation."

J.6. Summary Recommendations to LC. The MARC Pilot Project has been within its limited scope an unquestioned success. Effective use can indeed be made of catalogue data distributed centrally in machine-readable form. Our recommendation is that this distribution be continued. The fact that plans for MARC II are already well advanced suggests that this recommendation is neither inappropriate nor unrealistic.

Appendix B SPECIAL STUDIES

Monograph Acquisitions Patterns in Selected Libraries

A limited analysis of the monograph acquisition pattern of special technical libraries and public libraries was performed by Programming Services, Inc., to study the age and language of the materials received in each of these libraries or library systems. The language data would provide guidance to the Library of Congress for the future expansion of the MARC project. The age of materials would give an indication of the cumulation of retrospective bibliographic records necessary to satisfy the requirements of special technical libraries and public school libraries.

The age data is an approximation determined by subtracting the year of publication from the year of ordering.

The following libraries cooperated in the data collection for the study:

Public Libraries

1. Santa Clara Valley Library System (a countywide processing center for participating libraries), Santa Clara, California.
2. San Francisco Public Library (a large metropolitan library), San Francisco, California.
3. Palo Alto Public Library (a medium-sized city library), Palo Alto, California.
4. Morgan Hill Library (a small-town library), Morgan Hill, California.
5. Saratoga Public Library (a small-town library), Saratoga, California.

Special Technical Libraries

1. Stanford Research Institute Library, Menlo Park, California.
2. Varian Associates Technical Library, Palo Alto, California.
3. Redstone Scientific Information Center, Redstone Arsenal, Alabama.

The Santa Clara Valley Library data were taken from their order slips for 1,279 monographs requested during fiscal year 1965. This represented about 20 percent of all their monograph orders during that time period.

The Morgan Hill data were taken from their order slips for 505 monographs ordered during the fiscal year 1965. This represented an estimated 40 percent of the Morgan Hill monograph acquisitions during that period. Some of these Morgan Hill data are included in the data reported for the Santa Clara Valley System, of which Morgan Hill is a member. The Morgan Hill Library was in the process of rapid expansion into a new facility at the time that the analyzed orders were placed. This might account for its acquisition of a relatively large amount of older material.

The Saratoga data were taken from their order slips for 667 monographs requested during fiscal year 1965. This represented an estimated 40 percent of all their orders during this period. Saratoga Library represents a small local library that has been in operation for several years, serving a relatively stable population. Like Morgan Hill, Saratoga is a member of the Santa Clara Valley system.

The Palo Alto data were taken from their order slips for 1,155 monographs ordered during fiscal year 1966 and represent about 25 percent of their orders during that period. The data show fewer current acquisitions than the other libraries. It was later determined that this may have been because the library buys very little current fiction (contrary to the practice of many public libraries) but rents it instead. This library subscribes to the McNaughton Service, an arrangement by which the library rents a large number of copies of current fiction from the McNaughton Company and returns them when the demand drops. The library also has an option to buy the book at a later date. The point, however, is that the library avoids the necessity of cataloging all of the current and sometimes transient fiction works. San Francisco was the only other library in this study to subscribe to the McNaughton Service.

The San Francisco data were taken from order slips for 1,212 monographs ordered during the years 1964–65. This represents about nine percent of the monograph orders during a given year.

Because of the high volume of circulation of some public library materials, e.g., popular children's books, some books must be replaced with new ones as they wear out. Some of the orders included in the analyses of the libraries represent replacement rather than initial orders and thus bias the order data in the direction of older material. The first analysis of the order slips for these libraries did not differentiate between initial and replacement orders, consequently the mix of these order types is shown in Table 1. A subsequent analysis of the Saratoga, Santa Clara County, and San Francisco order data indicated that 15, 16, and 17 percent of the orders, respectively, were for replacements. A more detailed analysis showed that in all instances the replacement orders were for older material than the orders for initial copies.

The implications of the replacement orders for MARC are not clear. If the ordering library has a policy of not recataloging replacement orders, then no MARC data is needed for replacements, and the reported data reflect proportionally more older records than they should. But if the library has a policy of recataloging replacement books because of possible changes in the classification schedules or some other factor, then according to the composite age distribution given in Table 1, MARC records would be needed.

The Stanford Research Institute data was taken from the original order slips of 1,130 monographs requested during 1965. This was about half of the orders placed that year. The Varian data was taken from the order slips of all of the 1,038 monographs requested during the period 1961-October 1966. The Redstone data for 1,103 monographs was taken from a November 1966 report of their books currently on order and may be somewhat biased by possibly including many items (perhaps the older out-of-print items) that have accumulated on the listing because of delays in obtaining copies.

TABLE 1.—*Age Data*

Distribution of the Age of Monographs Acquired by Public Libraries Expressed in Percentages

	Just Published	Calendar Age of Materials Ordered in Years							
		1	2	3	4	5	6	7	Over 7
Santa Clara County Valley Library	52.0	23.7	3.0	1.9	1.6	1.6	2.3	.7	13.2
Morgan Hill Library	47.3	18.4	2.4	1.2	2.0	2.6	3.0	2.6	20.5
Palo Alto Public Library	39.4	22.1	5.7	4.6	3.1	3.4	2.5	2.2	17.0
San Francisco Public Library	43.3	22.8	5.5	3.5	4.9	4.1	2.5	1.8	10.6
Saratoga Public Library	63.4	16.0	2.7	.6	1.0	1.5	1.7	1.2	11.9

The above table suggests that about 80 percent of acquisitions could be covered by bibliographic records for the most recent seven years and that most public library acquisitions are for relatively current materials. The seven-year figure could cover an even higher percentage for special libraries.

TABLE 2.—*Age Data.*

Distribution of Age of Monographs Acquired by Special Libraries Expressed in Percentages

	Just Published	Calendar Age of Materials Ordered in Years							
		1	2	3	4	5	6	7	Over 7
Stanford Research Institute Library	36.4	32.5	12.2	5.2	3.3	2.1	1.7	1.2	5.4
Varian Associates Technical Library	56.6	20.0	6.2	3.8	3.4	1.8	2.1	.8	5.3
Redstone Scientific Information Center	31.8	24.4	11.5	8.4	5.1	2.6	4.5	.6	11.1

TABLE 3.—*Language Data*

	Percent of All Orders in Foreign Languages	Combined French and German	Percent of Foreign Language Orders in				
			French	German	Spanish	Italian	Other
Santa Clara Valley Library	1.0	42.9	28.6	14.3	14.3	38.1	4.7
Palo Alto Public Library	2.3	45.6	43.5	2.1	43.5	10.9	------
Morgan Hill Library	0.1	50.0	------	50.0	------	50.0	------
Saratoga Public Library	0.2	100.0	75.0	25.0	------	------	------
San Francisco Public Library	6.9	49.9	33.6	16.3	8.6	15.4	26.1

TABLE 4.—*Language Data*

	Percent of All Orders in Foreign Languages	Percent of Foreign Language Orders in French and German
Stanford Research Institute Library	6.1	49.2
Varian Associates Technical Library	7.5	80.8
Redstone Scientific Information Center	16.3	91.2

The above tables suggest that in any expansion to languages other than English, emphasis on French and German language cataloging would most satisfy the requirements of public and special libraries.

A similar study was performed by Programming Services, Inc., on the acquisition practices of the MARC participants. The following table is derived from a questionnaire submitted to the participants in the spring of 1967 and reflects the response of 14 out of the 16 participants.

TABLE 5.—*Monograph Acquisitions in 14 MARC Participant Libraries*

Name of Library or Institution	Distribution of Orders by Date of Imprint	Percent	Distribution by Language	Percent	Percent of Orders for Replacements or Added Copies
Argonne National Laboratory	Current year	73	English	88	1
	—1 year	14	German	5	
	—2 years	5	French	3	
	—3 years	1	Russian	3	
	—4 years	1	Other	1	
	—5 years +	6			
	(Includes added and replacement copies				
Georgia Institute of Technology	Current year	70	English	90	5
	—1 year	20	Russian	8	
	—earlier	10	German	1	
			Other	1	
	(Includes added and replacement copies)				
Nassau County Library	(No information)		English	99.9	50–65
National Agricultural Library	(No information)		English	50 (est.)	under 1
Redstone Scientific Information Center	Current year	20	English	90	30
	Past 6 years	60	German	8	
	Over 6 years	20	French	1	
			Russian	under 1	
			Italian	under 1	
Rice University	Current year—	60	English	65	1
	Past 1 year	10	German	20	
	Past 2–6 years	10	French	5	
	Over 6 years	20	Spanish	3	
			Russian	3	
			Italian	2	
			Other	2	
Washington State Library	Current year—	34	Almost All English		30
	Past 1–5 years	34			
	Past 6–10 years	9			
	Past 11–20 years	17			
	Over 20 years	6			

TABLE 5.—*Monograph Acquisitions in 14 MARC Participant Libraries*—Continued

Name of Library or Institution	Distribution of Orders by Date of Imprint	Percent	Distribution by Language	Percent	Percent of Orders for Replacements or Added Copies
UCLA	No figures are kept. About half of orders are for current materials.		English	50	3, replacements; 4, added copies
			Other	50	
University of Florida	1966	22	English	51	20 of in-print titles
	1965	19	Spanish & Portuguese	23	
	1964	10	French	11	
	1960–63	12	German	8	
	1956–59	6	Other	7	
	1950–55	6			
	1940–49	5			
	1930–39	10			
	pre-1930	10			
University of Toronto	(No information)		(No information)		30
Indiana University	1965 +	70	English	90	10–15
	1955–64	20	French	9	
	pre-1955	10	German or Spanish or Other	1	
Montgomery County Public Schools	1959–67	80	English	99.5	40–50
	1947–57	12	Spanish, French, German, Russian	0.5	
	pre-1947	8			
University of Chicago	1967	31	English	45–50	Less than 5
	1966	36	Foreign	50–55	
	1965	9	Non-Roman alphabet	10–15	
	1960–64	9			
	1940–59	5			
	pre-1940	10			
	(of that 10%, 6% is prior to 1900)				
University of Missouri	Past 5 years	90	(No information)		less than 10

Book-Oriented SDI Service Provided for 40 Faculty

William J. Studer

The following report was prepared by William J. Studer, doctoral candidate in the Graduate Library School at Indiana University. This study was conducted independently of the work done by Indiana University as a MARC participant. It was published in *Random Bits*, volume 3, numbers 3 and 4, and it is being reprinted in part with permission of the author and the publisher because it sheds valuable insight on the potential uses of machine-readable data.

In November 1966, the Library of Congress launched a pilot project in centralized production and distribution of its traditional and widely sold catalog records (3×5 printed cards) in machine-readable form. The basic aims of Project MARC (*MA*chine-*R*eadable *C*ataloging) are to test the feasibility and desirability of providing such a service and to gather data on local utilization and accompanying benefits and problems. The Indiana University Library is among 16 libraries chosen as official participants in the project. Each week the Library of Congress sends a magnetic tape with 600-800 new catalog records added to the master file.

Varying degrees of computer-based automation of academic library activities have been steadily gaining momentum over the past five years or so, but the majority of applications have been aimed at mechanization of internal library processes. Initial concern for internal efficiency is logically understandable, and, to be sure, improvement of internal management is of at least indirect benefit to users; but the average faculty user can see little evidence that the flurry of library automation is of significant benefit to him.

With this situation in mind, I decided to make experimental use of the Project MARC records to provide a service of direct use to individual faculty and which also should be indirectly useful to the library. The objective of this service is to furnish the faculty member regularly with computer-produced lists of new publications in his specified fields of interest. This automatic alerting system is most often referred to as Selective Dissemination of Information (SDI) or Current Awareness Service.

The first working SDI system was established in 1959 by IBM, and since then SDI systems have become almost epidemic and have gained widespread acceptance—mostly serving research personnel in the fields of science and technology for the purpose of providing manageable access to the flood of current periodical and technical report literature. However, applications of SDI outside the fields of science and technology have been very limited, and, to my knowledge, no one has attempted to provide such a service for book materials (books in the broad sense of nonperiodical, monographic publications).

It is my hypothesis that an SDI system concerned with book-type material would be of significant benefit to faculty in keeping them alerted to what is being published in their fields of interest—especially faculty in the nontechnical areas where books are probably still as vital, if not more important, a medium of information and ideas as periodical and report literature. And the volume to which book publishing has risen today also poses problems for the individual of identifying the small number of relevant items from among the whole.

The availability of machine-readable catalog records from the Library of Congress would seem to make such an SDI service possible. Our de facto national library collects and catalogs promptly a majority of the world's important literature, and therefore potentially provides as universal a base for the book-oriented SDI system as exists. The broad subject coverage of materials processed furnishes a single data base for creating lists of current publications for the many faculty who have cross-disciplinary interests. Most faculty lean heavily on several discipline-oriented, professional journals to inform them of new books in their fields of interest. Publications of potential and even direct interest often fall outside the province of these particular journals, and the individual is unable to scan all sources which might include notices and reviews of pertinent materials.

All computer-based SDI systems work on the same principle and include two basic elements: subject interest profiles for the users and a machine-readable file of indexed bibliographic records of current materials.

Interest profiles consist of legitimate index words (descriptors) which best describe a user's interests, and these profile descriptors are matched against the corresponding descriptors used to index publications in the file being searched. In essence, when a proper match occurs, the record for the particular item is printed out and sent to the user.

My efforts in setting up a pilot SDI system for the faculty at Indiana University began with the sending of invitations to participate in the experiment, hoping to interest 30-40 faculty in taking part. Letters were sent to 298 faculty members in seven areas of the social sciences (anthropology, business, economics, education, government, history, and sociology). Very unexpectedly, 209 expressed a desire to participate, which speaks rather positively for at least the initially felt need for such a service. Since I judged that I could not personally handle more than 40 users, this number was chosen from the positive replies on the basis of a random, proportional sample. The general area of social sciences was chosen somewhat arbitrarily because I felt best equipped to deal with faculty in this subject area. There is no sound reason to assume that the system would work less well for one subject area than another.

Most faculty members at I.U. had already created interest profiles as part of a form filled out for the Office of Research and Advanced Studies. These profiles were obtained and returned to the 40 participants with a request to review and update them for purposes of the SDI experiment. Interests were expressed in key words and phrases in the faculty members' own language. I translated these stated interests into the appropriate subject headings and subject classification numbers (hereafter called descriptors of terms) used by the Library of Congress to index its materials—keeping in mind the fact that publications are indexed rather shallowly, averaging somewhat under three descriptors per item. Suffice it to say that the translation of broad and narrow interests, expressed in uncontrolled natural language, into an equivalent set of descriptors chosen from controlled authority lists (embracing the universe of knowledge and including tens of thousands of descriptors) is a difficult task and requires some ingenuity. However, a profile is the key element in an SDI system. The quantity and quality of the output from the system is directly proportional to the quality of the profile; so the job must be done as carefully and accurately as possible. Profile terms are stored on punched cards for computer processing.

Records on the MARC master tape sent weekly from the Library of Congress are tagged to make current additions identifiable (N = added to the tape this week; L = added to the tape last week). The experimental SDI system runs on a two-week cycle, and thus every second MARC tape is used to create a current search file. Three computer programs (written in FORTRAN 63) and three processing steps are involved in producing the SDI lists, which are accomplished on the CDC 3400/3600 system at I.U.'s Research Computing Center:

1. The first process uses the MARC master tape as input. The program extracts records tagged with N or L and copies these records onto a new tape (New MARC file) for subsequent printout. As part of the same processing, the descriptors and unique accession number of each record are extracted and copied onto another tape (MARC extract) for subsequent matching of profile terms.

2. The second process uses the MARC extract tape and profile descriptors as input, and the program executes the matching of descriptors. When a proper match occurs, the accession number for the item is recorded, together with a profile identifier on another tape (Accession numbers).

3. The final process uses the accession numbers tape and new MARC file tape as input, and the program matches on accession numbers. When a match occurs, the pertinent bibliographic record is printed out in triplicate. The individual records are then assembled into three identical lists for each user. One copy is for his own use; the records on another are evaluated (according to codes appearing as the last line) and the copy returned as feedback to the SDI system; and the third copy is used, if desired, for recommending items in the list

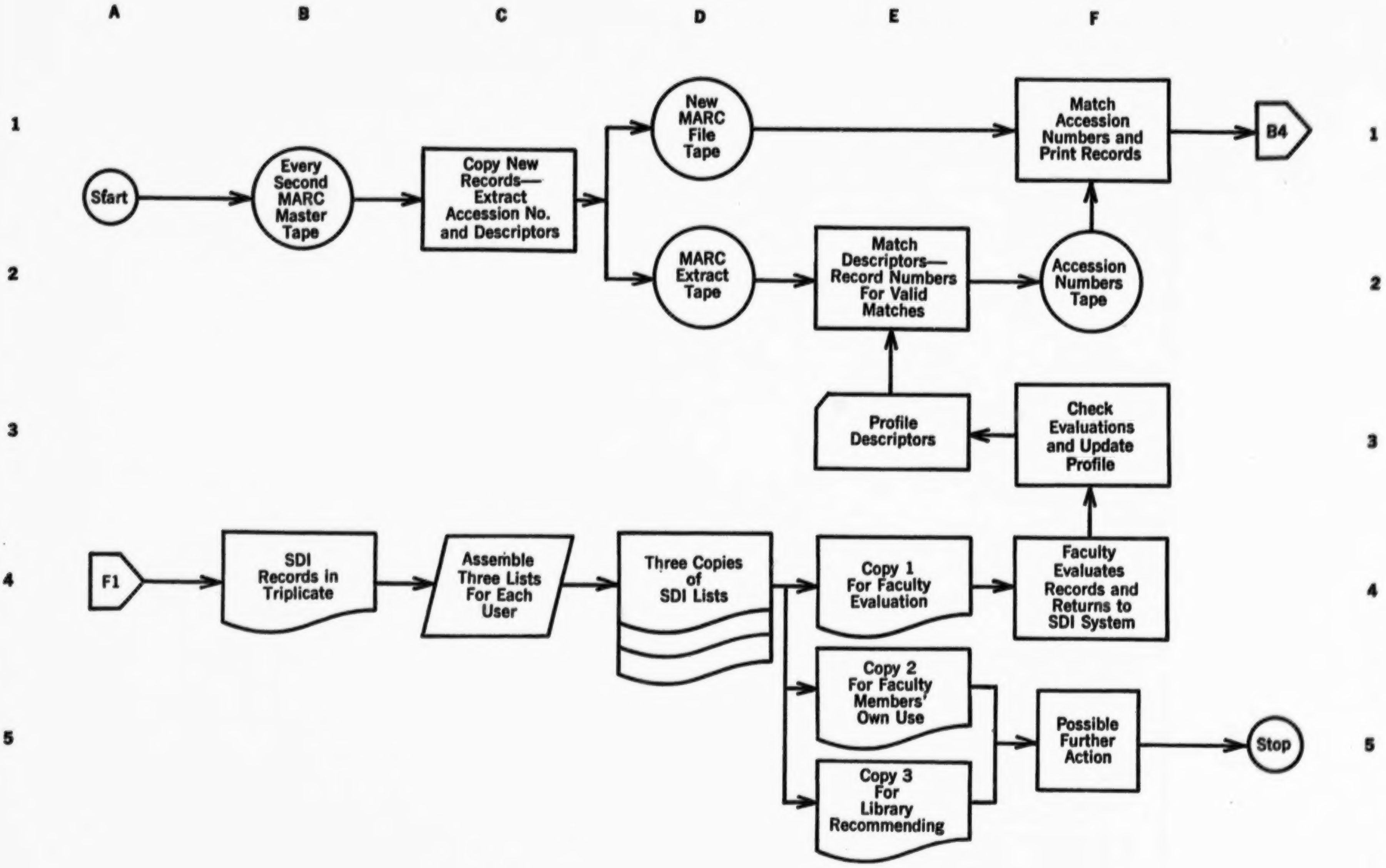
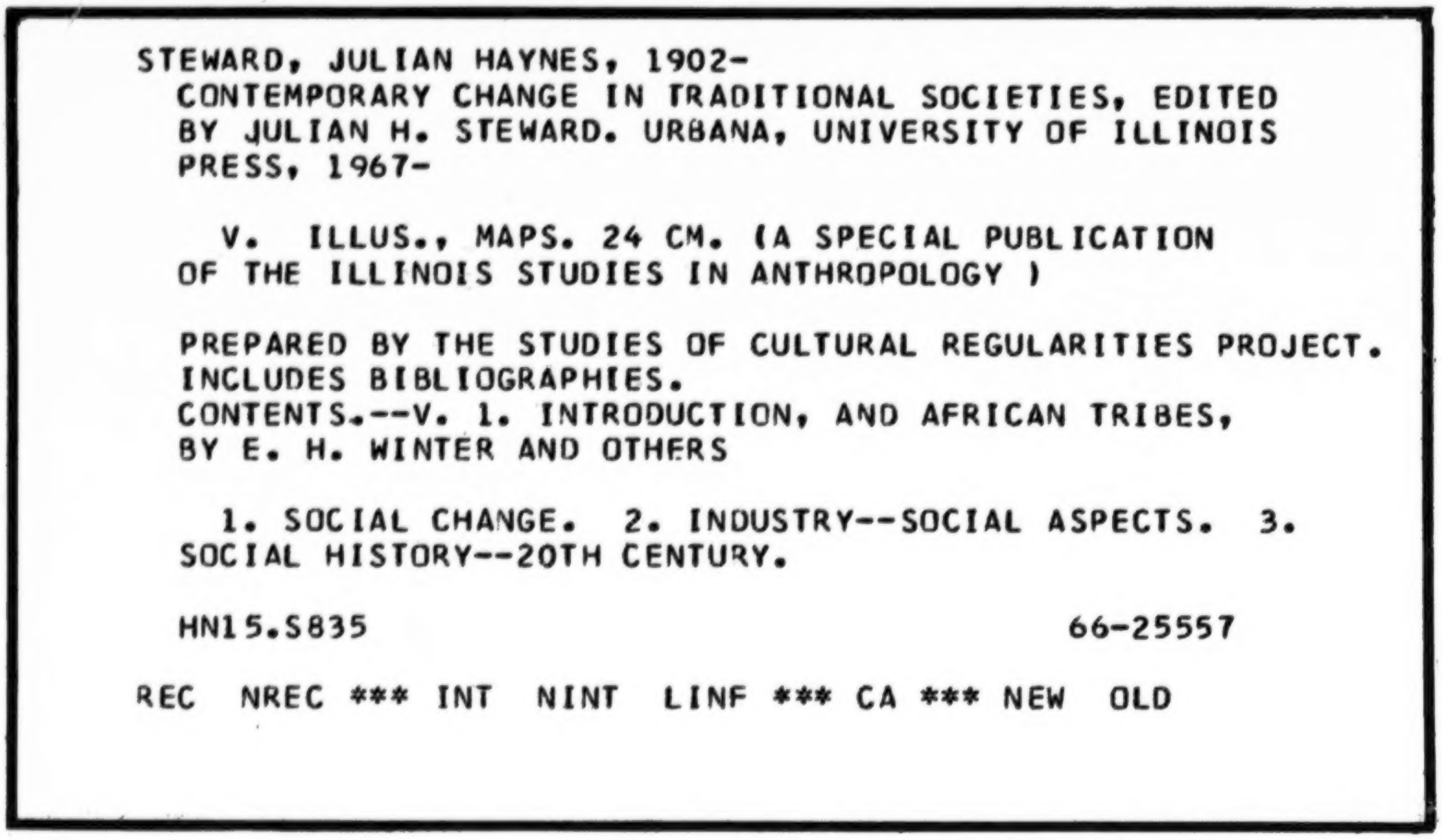

A
B
C
D
E
F
1
2
3
4
5
Start
Every Second MARC Master Tape
Copy New Records— Extract Accession No. and Descriptors
New MARC File Tape
MARC Extract Tape
Match Accession Numbers and Print Records
B4
Match Descriptors— Record Numbers For Valid Matches
Accession Numbers Tape
Profile Descriptors
Check Evaluations and Update Profile
F1
SDI Records in Triplicate
Assemble Three Lists For Each User
Three Copies of SDI Lists
Copy 1 For Faculty Evaluation
Faculty Evaluates Records and Returns to SDI System
Copy 2 For Faculty Members' Own Use
Copy 3 For Library Recommending
Possible Further Action
Stop

FIGURE 1.—*SDI System Using MARC Records from the LC*

for library acquisition. A block flow diagram of the system operation and a sample record printout are presented in Figures 1 and 2, respectively. The record includes basically the same data as would be found on a printed library catalog card.



```
STEWARD, JULIAN HAYNES, 1902-
  CONTEMPORARY CHANGE IN TRADITIONAL SOCIETIES, EDITED
  BY JULIAN H. STEWARD. URBANA, UNIVERSITY OF ILLINOIS
  PRESS, 1967-

    V.  ILLUS., MAPS. 24 CM. (A SPECIAL PUBLICATION
  OF THE ILLINOIS STUDIES IN ANTHROPOLOGY )

  PREPARED BY THE STUDIES OF CULTURAL REGULARITIES PROJECT.
  INCLUDES BIBLIOGRAPHIES.
  CONTENTS.--V. 1. INTRODUCTION, AND AFRICAN TRIBES,
  BY E. H. WINTER AND OTHERS

    1. SOCIAL CHANGE.  2. INDUSTRY--SOCIAL ASPECTS.  3.
  SOCIAL HISTORY--20TH CENTURY.

  HN15.S835                                  66-25557

REC  NREC *** INT  NINT  LINF *** CA *** NEW  OLD
```

FIGURE 2.—*Sample Record Printout*

The user evaluates records: (1) by indicating that the publication cited is of interest (INT); is not of interest (NINT); or that the citation lacks sufficient information for him to judge the item's relevance (LINF); and (2) by indicating whether the publication is new to him (NEW) or is already known to him (OLD). The first evaluation tells me how well the system performs in selecting relevant publications, and adjustments in profiles are made according to this response. The second evaluation codes give information on whether the faculty member intends to recommend a given title for library acquisition (REC = yes; NREC = no) and intends to call the publication to the attention of a colleague (CA).

As of this writing, only two lists have been produced and circulated, but initial feedback would seem to indicate that the system is operating successfully. On the average, well over half of the citations are judged to be relevant items, and over 75 percent are new to the users. Also, a great majority of faculty indicate that they are using the lists as a convenient means of recommending books for library acquisition, whereas previous to the receipt of these lists most said that they recommended only occasionally or seldom. Generally speaking, in an academic library setting, faculty, as subject specialists, are supposed to be the prime movers in selecting materials in their fields for the library's collections. In fact, however, most faculty leave this essential task to the librarians. If SDI lists do indeed stimulate active participation in recommending, the library can benefit in two ways: (1) improvement in the quality and coverage of its collections; and (2) the facilitating of the acquisition process because records from the SDI lists include all the correct bibliographic information needed for ordering.

A weighted-term search strategy is used in computer matching of profile terms against the record descriptors, both of which are limited to 16 characters in length. Profile tehms may be assigned weights between +9 and −9, and each profile is assigned a cutoff weight, thus simulating the AND, OR, NOT relationships of Boolean algebra. These relative weights determine whether a given record will register as a valid match. For example, if a cutoff weight of 6 is assigned, the profile descriptors which match record descriptors must carry weights which total at least 6—also taking into account the possible presence of matching terms with negative weights.

The computer program being used for matching was written by the Aerospace Research Applications Center (ARAC) at I.U. for use in its own information retrieval services, and it was altered to fit some special requirements of my SDI system. All of the programming for the SDI system has been accomplished under the direction of Richard W. Counts, ARAC's Assistant Director for Information Systems.

In an interview held with each participant prior to beginning the SDI service, most stated that they made a direct effort to learn of newly published books in their fields but admitted that their present techniques and existing sources for learning of new materials do not accomplish the job optimally. Also, most agreed that it is either vital or important to learn of new works as soon as they are published. Insufficient time, scattering of pertinent items in multiple sources, and time lag between publication date and listing of the book in some source were frequently cited as problems. Hopefully, SDI lists generated from a file potentially as catholic and voluminous as Library of Congress current acquisitions can provide a single, compact source for promptly alerting faculty to new publications spanning their various fields of interest.

Presently only a selection of English language publications are entered on the MARC master tapes sent from the Library of Congress and these records represent but a fraction of the Library's actual intake. But Project MARC is no longer considered a pilot project. It has proved successful, and the Library hopes to be able to provide the service commercially by July 1968. The system will be expanded to include virtually all English language materials, and gradually various foreign language materials will be added—thus greatly increasing the comprehensiveness of the data base.

The SDI system described here does indeed work, but it is only a pilot model of a potential final system. Since this is a dissertation project, I have developed and must operate the system with only my own resources. For the sake of workability and time limitations, I have had to make compromises and adjustments that would not be necessary with adequate financial support, administrative authority and control, a full-time staff, and tailormade programming. However, my purpose is to test the basic feasibility and usefulness of book-oriented SDI service for faculty. Many refinements can and would certainly be made for a permanent installation. I believe that the idea is sound and that the extension of SDI service to all faculty would be economically viable if an academic institution is as willing to fund this kind of service as the many other services with which it now provides its faculty. However, reasonable judgment as to the practicality and relative benefits of faculty SDI service must await the participants' overall reaction during the test period and a final assessment of costs and a host of other factors.

☆U.S. GOVERNMENT PRINTING OFFICE: 1969 O—312-671